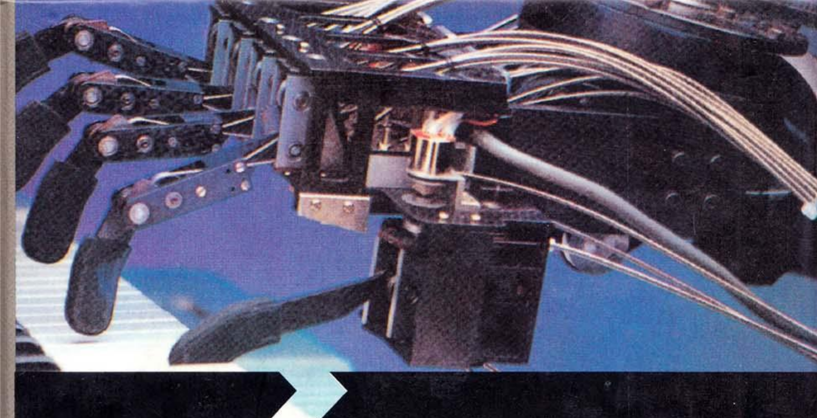


El hombre mecánico

El futuro de la robótica y la inteligencia humana

Hans Moravec

Biblioteca
Científica
Salvat

El hombre mecánico

El futuro de la robótica y la inteligencia humana

Hans Moravec

SALVAT

Versión española de la obra *Mind Children. The Future of Robot, and*
Human Intelligence, publicada de acuerdo con Harvard University Press

Traducción: Ana Mendoza
Diseño de cubierta: Ferran Cartes / Montse Plass

Edición digital: Sargont (2016)

ISBN: 84-345-8880-3 (Obra completa)
ISBN: 84-345-8891-9 (Volumen 11)
Depósito Legal: B-27525-1993
Publicada por Salvat Editores, S.A., Barcelona
Impresa por Printer, i.g.s.a., Octubre 1993
Printed in Spain

INDICE

A la memoria de mi padre,
que me enseñó a jugar con máquinas.
A mi madre,
que me enseñó a leer.
A Ella,
que me completó.

PROLOGO

Nuestros genes han estado sumergidos durante miles de millones de años en la espiral de una carrera de armamentos inexorable, luchando unos contra otros, pero finalmente han conseguido su objetivo. Han construido un arma tan poderosa que, con ella, podrán derrotar tanto a los vencedores como a los vencidos. Y no es la bomba de hidrógeno. Con el empleo generalizado de las armas nucleares lo único que se conseguiría es retrasar el final, mucho más interesante, que se ha diseñado.

Lo que nos espera no es el olvido, sino un futuro que, desde nuestra ventajosa situación actual, se puede describir con las palabras *«posbiológico»* o, mejor aún, *«sobrenatural»*. En ese mundo, la marea del cambio cultural ha barrido al género humano y lo ha sustituido por su progenie artificial. Las consecuencias finales todavía nos resultan desconocidas, aunque muchos de los pasos intermedios no sólo son predecibles sino que ya se han dado. En la actualidad, nuestras máquinas son todavía creaciones simples que precisan los cuidados maternales y la atención constante de todos los recién nacidos. Y no se puede decir que merezcan el calificativo de *«inteligentes»*. Pero en el curso de un siglo madurarán, se convertirán en seres tan complejos como nosotros y, finalmente, en entes que trascenderán todo lo que conocemos, de los que nos podremos sentir orgullosos y considerarlos nuestros propios descendientes.

Estos frutos de nuestra mente, liberados del laborioso ritmo de la evolución biológica, serán libres para crecer y enfrentarse a los desafíos, extraordinarios y fundamentales, del universo. Nosotros, los seres humanos, nos beneficiaremos de sus logros durante un cierto tiempo, pero antes o después, como nuestros hijos naturales, buscarán su propio camino mientras que nosotros, sus ancianos progenitores, desapareceremos silenciosamente. No se perderá mucho cuando les pasemos el relevo. Nuestra descendencia artificial, por su bien, recordará casi todo

lo relacionado con nosotros, incluso los logros detallados de las mentes humanas independientes.

El proceso comenzó hace unos cien millones de años, cuando ciertas familias de genes encontraron la forma de crear animales con capacidad para aprender ciertos comportamientos de sus progenitores mientras éstos estaban vivos, en vez de heredarlos en el momento de la concepción. Se intensificó hace diez millones de años, cuando nuestros antepasados primates empezaron a utilizar herramientas hechas de huesos, palos y piedras, y experimentó una nueva aceleración con la utilización del fuego y el desarrollo de los lenguajes complejos, hace un millón de años aproximadamente. Para la fecha en que apareció nuestra especie —hace unos cien mil años—, la evolución cultural, el coloso que inconscientemente habían construido nuestros genes, avanzaba con un ímpetu irresistible.

Durante los últimos diez mil años, los cambios experimentados por los genes han sido intrascendentes en comparación con los tremendos y rápidos avances de la cultura humana. Hemos sido testigos, en primer lugar, de una revolución agrícola a la que ha seguido el establecimiento, a gran escala, de gobiernos burocráticos con poder para recaudar impuestos para su propio mantenimiento, el desarrollo de las lenguas escritas y la aparición de clases acomodadas con tiempo y energía para dedicarse a temas intelectuales. Aproximadamente en los últimos mil años, algunas creaciones, empezando por la imprenta, han acelerado enormemente el movimiento de la información cultural y, en consecuencia, su ritmo evolutivo.

Con la aparición en escena de la revolución industrial, hace doscientos años, entramos en la fase final. Los sustitutos artificiales de las funciones corporales humanas como, por ejemplo, la de levantar algo o transportarlo, se han ido haciendo cada vez más interesantes, económicamente hablando. Mejor dicho, indispensables. Posteriormente, hace cien años, con la invención de las calculadoras mecánicas, hemos podido duplicar, por vez primera, algunas funciones de la mente humana, triviales pero fastidiosas.

Desde entonces, la potencia de cálculo de los dispositivos mecánicos se ha multiplicado por mil cada veinte años.

Nos aproximamos al momento en que prácticamente todas las funciones humanas esenciales, físicas y mentales, tendrán su equivalente artificial. El prototipo de la unión de todos los resultados del desarrollo cultural será el robot inteligente, una máquina que podrá pensar y actuar como un ser humano, aunque física y mentalmente no lo sea. Estas máquinas podrán llevar a cabo nuestra evolución cultural, incluyendo su propia construcción, y perfeccionarse cada vez más sin nosotros y sin nuestros genes. Cuando eso suceda, nuestro ADN se encontrará sin trabajo y, además, habrá perdido la carrera evolutiva en un nuevo tipo de competición.

A. G. Cairns-Smith, un químico que ha contemplado los principios de la vida en la tierra primitiva, denomina a esta clase de golpe interno *relevo genético.* Y sugiere que ha sucedido anteriormente, por lo menos en una ocasión. En su obra *Seven Clues to the Origin of Life,* Cairns-Smith afirma que los precursores de la vida, tal y como la conocemos, fueron cristales microscópicos de arcilla que se reproducían por el simple proceso de la expansión cristalina. La mayor parte de estos cristales están alterados por modelos de dislocación en la disposición ordenada de sus átomos, muchos de los cuales se propagan a medida que crece el cristal. Si se rompe el cristal, cada pedazo hereda una copia del modelo, a veces con una ligera modificación. Estos defectos afectan muchísimo a las propiedades físicas y químicas de la arcilla. Los cristales con un determinado modelo de dislocación pueden formar grupos muy densos, mientras que los de otro modelo pueden agregarse y constituir masas esponjosas. Las aguas que llevan minerales se pueden apartar de uno de estos tipos y acercarse a otro, proporcionando materias primas para que siga adelante el crecimiento.

Estos modelos también afectan indirectamente al crecimiento, ya que modifican la composición química de otras moléculas del entorno. Las arcillas son catalizadores químicos fuertes. Los diminutos cristales tienen una enorme área superficial total a la que se pueden adherir las moléculas de ciertas

configuraciones, dependiendo de la forma externa del cristal y de la molécula en cuestión. Es decir, los cristales corrientes poseen los factores indispensables de la evolución darwiniana: reproducción, herencia, mutación y diferencias en el éxito de la reproducción.

De acuerdo con la teoría de Cairns-Smith, el primer relevo genético comenzó cuando algunas familias de arcilla, en fuerte competencia darwiniana con otras, empezaron a codificar externamente algunas informaciones genéticas en las largas moléculas de carbono. Estos polímeros son mucho más estables que los modelos de dislocación, fácilmente perturbables, y los organismos que más los utilizan tienen más éxito en la reproducción. Aunque en un principio dependen totalmente de los mecanismos químicos existentes, asentados en el cristal, a medida que cobra más importancia su papel reproductor, van dejando de estar supeditados a los cristales. Con el tiempo, desaparece el andamiaje simple del cristal y deja tras de sí, en su camino evolutivo, el sistema interdependiente de maquinaria orgánica que denominamos vida.

Hoy, miles de millones de años después, se está produciendo otro cambio en la forma en que se transmite la información de una generación a otra. Los seres humanos hemos evolucionado a partir de organismos definidos casi en su totalidad por sus genes orgánicos. En la actualidad, contamos también con una vasta colección de informaciones culturales, que aumenta rápidamente, generada y almacenada más allá de nuestros genes: en nuestro sistema nervioso, en las bibliotecas y, más recientemente, en los ordenadores. Nuestra cultura sigue dependiendo totalmente de los seres humanos biológicos, pero cada año que pasa nuestras máquinas, uno de los productos fundamentales de la cultura, asumen un papel más importante tanto en la conservación de la misma como en su continuo desarrollo. Antes o después, nuestras máquinas llegarán a tal grado de erudición que podrán dedicarse a su conservación, reproducción y perfeccionamiento sin necesidad de ayuda. Cuando suceda, se habrá terminado el nuevo *relevo genético*.

Nuestra cultura podrá evolucionar con independencia de la biología humana y de sus limitaciones, y pasará de una generación de máquinas capaces e inteligentes a otra mucho más perfecta.

Nuestros genes biológicos, y los cuerpos de carne y hueso que han creado, tendrán un papel cada vez menos importante en este nuevo sistema. Pero ¿acaso nuestras mentes, donde nació la cultura, se perderán? Es posible que no. La futura revolución puede liberar a las mentes humanas con la misma eficacia que libera a la cultura humana. En la situación actual somos híbridos incómodos de la biología y la cultura, y muchas de nuestras características biológicas se han quedado rezagadas respecto de las creaciones de nuestra mente. Nuestras mentes y nuestros genes pueden tener muchos objetivos en común durante la vida, pero existe una tensión entre el tiempo y la energía que se utilizan para crear, desarrollar y difundir las ideas y el esfuerzo que se realiza para producir una nueva generación (como pueden corroborar todos los padres con hijos adolescentes). La difícil tregua que se establece entre la mente y el cuerpo se rompe por completo cuando termina la vida. Nuestros genes suelen sobrevivir a nuestra muerte agrupados de distintas maneras, bien en nuestros hijos, bien en nuestros parientes. De una forma muy sutil, no cabe ninguna duda sobre su interés evolutivo en realizar experimentos como éste después de barajar los naipes genéticos. Pero el proceso resulta demoledor para nuestra otra mitad. Porque, simplemente, perecen con nosotros muchísimos aspectos de nuestra existencia mental que habíamos conseguido tras arduos esfuerzos.

No es difícil de imaginar al pensamiento humano liberado de la esclavitud del cuerpo mortal: la creencia en otra vida después de la muerte está ampliamente extendida. Pero no es necesario adoptar ninguna postura ni mística ni religiosa para asumir esta posibilidad. Los ordenadores proporcionan un modelo que le resultará válido hasta al mecanicista más ardiente. Un cálculo que se está llevando a cabo, que podemos denominar razonablemente «proceso de pensamiento» del ordenador, se puede detener a la mitad y transferir, en forma de programa

y de datos sacados de la memoria de la máquina, a otro ordenador físicamente diferente y continuar trabajando en él como si no hubiera pasado nada. Imaginemos que la mente humana se pudiera liberar del cerebro de una forma parecida (aunque mucho más fascinante, técnicamente hablando).

Para que la mente pudiera funcionar con efectividad, necesitaría muchas modificaciones, después de que se la liberara de las limitaciones de un cuerpo mortal. La mentalidad humana natural está organizada para una progresión que abarque el período de una vida, desde la impresionable plasticidad hasta la rigidez. Por lo tanto, no es un material muy adecuado para la inmortalidad. Para que fuera viable, habría que reprogramarla con objeto de que se fuera adaptando continuamente. Mientras que un organismo mortal pasajero puede dejar la tarea de adaptación a los procesos externos de mutación y selección natural, una mente que aspire a la inmortalidad, tenga sus orígenes en un ser humano mortal o se trate de una creación artificial, debe estar preparada para adaptarse constantemente desde el interior.

A lo mejor tendría que someterse a un rejuvenecimiento cíclico y adquirir *hardware* y *software* nuevos en fases periódicas semejantes a la niñez. O acaso debiera actualizar continuamente el contenido de su mente y de su cuerpo, añadiendo, borrando y comprobando distintos componentes en todo tipo de combinaciones para mantenerse a la altura de las condiciones cambiantes. La comprobación tiene una importancia clave: dirige la evolución. Si el individuo toma demasiadas decisiones erróneas en estas comprobaciones, es que ha fracasado por completo, según la manera darwiniana antigua.

Un mundo posbiológico, dominado por máquinas pensantes que se perfeccionan a sí mismas, sería tan distinto de nuestro mundo actual de seres vivos como lo es nuestro mundo de la química inerte que le precedió. Resulta bastante difícil imaginar que sus habitantes fueran los frutos de nuestra mente. De todas maneras, vamos a intentar fantasear sobre algunas de las posibles consecuencias.

I. LA MENTE EN MOVIMIENTO

Tengo la absoluta seguridad de que los robots con inteligencia humana serán algo corriente dentro de cincuenta años. Si las comparamos con ellos, las máquinas más perfectas de la actualidad serán como la mente de los insectos frente a la de los seres humanos. Y esto, que tendrá lugar en las próximas décadas, será un gigantesco paso adelante.

Las imitaciones mecánicas de ciertas funciones humanas existen desde hace siglos. Muchos de los relojes de torre de la Edad Media cuentan con mecanismos que marcan las horas por medio de complicadas representaciones humanas de la moralidad, como santos, caballeros, obispos, ángeles, demonios y todo tipo de animales. Y, por lo menos desde el siglo XV, han servido de entretenimiento a la buena sociedad otros dispositivos más pequeños que caminaban, hablaban, nadaban, respiraban, comían y escribían con plumas de ave. Leonardo da Vinci construyó para sus mecenas complejos dispositivos mecánicos como los ya mencionados. Estas primitivas máquinas para medir el tiempo —que se movían impulsadas por agua corriente, pesos que caían o resortes— imitaban los movimientos de las cosas vivas, pero no eran sensibles al mundo que las rodeaba. Solamente podían *actuar*, aunque fuera de un modo encantador.

Las tecnologías eléctrica, electrónica y de radio, que se desarrollaron a principios de este siglo, permitieron la construcción de máquinas capaces de *reaccionar* ante la luz, el sonido y el control remoto invisible. El resultado fueron los entretenidos robots de demostración y un montón de conjeturas e historias sobre las máquinas del futuro, semejantes a los seres humanos. Pero al principio sólo se podían hacer conexiones simples entre sensores y motores. Estas máquinas más modernas pueden sentir y actuar, pero no *pensar*.

MÁQUINAS QUE PIENSAN (LIGERAMENTE)

En el curso de la Segunda Guerra Mundial se diseñaron ordenadores analógicos —máquinas que simulaban sistemas físicos por medio de una analogía entre sus cantidades variables y ciertas variaciones de ejes o voltajes— para controlar los cañones antiaéreos, la navegación y los bombardeos de precisión. Algunos investigadores observaron que existía una semejanza entre el funcionamiento de estos dispositivos y los sistemas reguladores de los seres vivos, y se inspiraron en esta semejanza para construir máquinas que actuaban y pensaban como si estuvieran vivas. Norbert Wiener, del Massachusetts Institute of Technology (MIT) acuñó el término *cibernética* para designar el estudio unificado del control y la comunicación en los animales y las máquinas. Los que la utilizaron combinaron esta nueva teoría sobre la regulación de la realimentación *(feedback)* con los avances de posguerra en el campo de la electrónica y con los primeros conocimientos sobre los sistemas nerviosos de los seres vivos, y construyeron máquinas que eran capaces de responder como animales simples y de aprender. Eran los rudimentos del pensamiento artificial.

Entre los logros más notables de la cibernética mencionaremos la serie de tortugas electrónicas que construyó en los años cincuenta un psicólogo británico, W. Grey Walter. Las primeras versiones llevaban cerebros dotados de válvulas electrónicas en miniatura, fotoválvulas como ojos, micrófonos como orejas, y unas varillas de contacto a modo de sensores. Cuando las pilas se les estaban terminando, podían localizar el punto en el que las recargaban, con lo que se evitaba el problema de que anduvieran moviéndose sin rumbo fijo. Algunas de ellas exhibían comportamientos sociales complejos cuando respondían a las luces de control y a los contactos con otras. Posteriormente, se acondicionó una de estas máquinas para que asociara un cierto estímulo con otro y, repitiendo una experiencia, fuera capaz de aprender, por ejemplo, que después de un cierto ruido fuerte siempre recibía un puntapié en el caparazón.

Una vez educada, la tortuga evitaba el ruido de la misma manera que anteriormente había reaccionado ante el golpe. Las asociaciones se iban acumulando lentamente, en forma de cargas eléctricas, en unos dispositivos denominados *capacitores*, que hacían el papel de memoria.

Es posible que la creación más impresionante de los cibernéticos haya sido la Bestia de John Hopkins. La construyó un grupo de investigadores del cerebro a principios de la década de los sesenta. Vagaba por las salas guiada por sonar, y con su ojo, que era una célula fotoeléctrica especial, buscaba la placa negra característica de los enchufes. Por último, se conectaba ella misma a la red. La Bestia sirvió de inspiración a muchos imitadores. En algunas copias se emplearon circuitos especiales conectados a cámaras de televisión en vez de células fotoeléctricas. La función de control se llevaba a cabo por medio de montajes de las entonces nuevas puertas lógicas digitales a transistores, que eran como las que se encuentran en la actualidad, a miles y a millones, en los circuitos integrados de los ordenadores. Algunas de ellas añadían movimientos nuevos al repertorio clásico, como, por ejemplo, «temblar para desanudar el brazo recargador».

La cibernética se desarrolló durante menos de dos décadas. Como sucede con frecuencia, la eclipsó un pariente suyo: la informática. Los pequeños ordenadores analógicos que inspiraron la cibernética durante la guerra tenían primos digitales mucho mayores. Estas máquinas no realizaban sus cálculos midiendo giros de ejes ni flujos de corriente, sino computando saltos discretos. Los primeros ordenadores digitales automáticos —calculadoras autónomas inmensas e inmóviles— se terminaron a finales de la guerra. El Colossus, una máquina británica ultrasecreta que descifró el código alemán Enigma y ayudó a modificar el curso de la guerra, registraba claves de códigos a una velocidad decenas de miles de veces superior a la de cualquier ser humano. En los Estados Unidos, ENIAC calculó tablas de artillería antiaérea para el ejército y, posteriormente, realizó operaciones para la construcción de la bomba atómica a una velocidad semejante a la del Colossus.

En un campo menos beligerante, estos «cerebros gigantes», como se les ha denominado, proporcionaron oportunidades sin precedentes para realizar experimentos muy complejos. Algunos pioneros, como Alan Turing, uno de los creadores del Colossus, y John von Neumann, que participó en la construcción de las primeras máquinas americanas, abrigaban la esperanza de que se pudiera encerrar en una máquina la capacidad de pensar racionalmente, nuestra única ventaja para enfrentarnos con el mundo. Los ordenadores pueden ensanchar nuestra mente, de la misma manera que las máquinas de vapor de la revolución industrial incrementaron la potencia de nuestros músculos. En 1950 Claude Shannon, del MIT, y Turing diseñaron programas que permitían a la máquina razonar y jugar al ajedrez y a otros juegos intelectuales, pero los ordenadores primitivos eran demasiado limitados y costosos como para darles este uso. A principios de la década de los cincuenta aparecieron algunos programas de juego de damas, bastante malos, para las primeras máquinas comerciales, y a finales de la década surgieron programas de ajedrez, igualmente malos, y un buen programa de damas diseñado por Arthur Samuel, de IBM. En 1957, Allen Newell y Herbert Simon, de Carnegie Tech (Carnegie Mellon University en la actualidad), y John Shaw, de RAND Corporation, presentaron el Teórico Lógico, el primer programa capaz de razonar sobre temas arbitrarios, que empezaba por axiomas y aplicaba reglas de inferencia para demostrar teoremas.

En 1960, John McCarthy, que por entonces estaba en el MIT, acuñó el término *inteligencia artificial* (IA) para definir los métodos capaces de hacer pensar a los ordenadores. En 1965, los primeros estudiantes de McCarthy, Marvin Minsky (también del MIT), Newell y Simon habían creado programas de IA que demostraban teoremas de geometría, resolvían problemas tomados de *tests* de inteligencia, libros de álgebra y exámenes de cálculo, y jugaban al ajedrez al mismo nivel que un estudiante universitario medio de primer año. Cada uno de los programas podía dedicarse solamente a un tipo limitado de problemas, pero, por vez primera, los resultados eran alentadores. Tanto, que la gente que trabajaba en este campo llegó a

pensar que, con una década más de progresos, se conseguiría una máquina auténticamente inteligente. Es comprensible este error de cálculo.

Hoy, un cuarto de siglo después, los ordenadores son miles de veces más potentes que estos modelos de los años sesenta, pero no parecen mucho más listos. Hacia 1975, los progresos en el campo de la inteligencia artificial, que hasta entonces habían sido como la carrera veloz de un puñado de entusiastas, se convirtieron en la laboriosa caminata de una multitud de trabajadores. Pero, aun así, los modestos éxitos que se han logrado contribuyen a mantener viva la esperanza. Los denominados *sistemas expertos*, es decir, los programas en los que se codifican las reglas de decisión de personas expertas en temas tales como diagnóstico de enfermedades, horarios laborales de fábricas o configuración de sistemas de ordenadores, tienen un lugar cada vez más importante en el mundo de los negocios. En el MIT, después de una labor de quince años, se reunieron en un programa maravilloso conocimientos de álgebra, cálculo, trigonometría y otros campos afines. Este programa, el MACSYMA, actualmente comercializado, maneja fórmulas simbólicas y ayuda a resolver problemas que, en otro caso, no se podrían solucionar. Algunos programas de ajedrez han sido calificados oficialmente para nivel de maestro, y en otros juegos, como el backgammon, se han conseguido unos resultados excelentes. Existen algunos programas semiinteligentes que entienden de ciertos temas restringidos, redactados en un inglés simple y a máquina, y hacen deducciones elementales mientras responden a preguntas relacionadas con esos textos. Algunos interpretan órdenes orales compuestas de palabras tomadas de un grupo de mil, previamente fijado, y otros pueden llevar a cabo tareas visuales sencillas como, por ejemplo, decidir si una pieza está colocada en su sitio.

Por desgracia para los robots humanoides, lo que peor hacen los ordenadores son las cosas más naturales para los seres humanos: ver, oír, manipular objetos, aprender idiomas y razonar con sentido común. Esta dicotomía —las máquinas hacen

bien cosas que a los humanos nos resultan muy difíciles, mientras que hacen mal las cosas que nos resultan más fáciles— es una ayuda enorme para resolver el problema de construir una máquina inteligente.

MÁQUINAS QUE VEN (DÉBILMENTE) Y AGARRAN (TORPEMENTE)

A mediados de los sesenta, los estudiantes de Marvin Minsky, en el MIT, empezaron a conectar a sus ordenadores cámaras de televisión y brazos mecánicos de robot, es decir, concedieron ojos y manos a las mentes artificiales para que sus máquinas pudieran ver, planificar y actuar. En 1965, estos investigadores habían creado una máquina capaz de encontrar y retirar bloques blancos de la superficie de una mesa negra. Para conseguirlo, hubo que elaborar un programa de control tan complejo como cualquiera de los programas de razonamiento puro que se empleaban en esa época, programas que, sin extremidades de robot, resolvían problemas de cálculo con la misma facilidad que un estudiante universitario de primer año. Sin embargo, el sistema ojo-mano de Minsky se podía mejorar. Los experimentos siguieron adelante no sólo en el MIT, sino también en otros lugares y, gradualmente, se fue creando un campo que hoy en día recibe el nombre de *robótica*, término que aparece por vez primera en 1942, en un relato de ciencia ficción de Isaac Asimov, y se deriva de la palabra *robot*, que inventó en 1921 el dramaturgo checo Karel Capek. La robótica comenzó en un punto mucho más bajo, en la escala del rendimiento humano, que la inteligencia artificial, pero sus progresos en los últimos veinte años han sido igual de angustiosamente lentos y dificultosos.

Ni todos los robots ni todas las personas se pasan la vida en la universidad. Hay muchos que trabajan para ganarse el sustento. Antes incluso de la revolución industrial, antes de que se mecanizara ningún pensamiento, ya existía la maquinaria parcialmente automática que, impulsada por el viento o por un

flujo de agua, molía grano o cortaba leña. El comienzo de la revolución industrial, en el siglo XVIII, lo marca la invención de una plétora de dispositivos que sustituían el trabajo manual de una manera precisa y sin participación humana. Estas máquinas, movidas por ejes que hacía girar el agua o el vapor, bombeaban, molían, cortaban, hilaban, tejían, imprimían, movían materiales y piezas y hacían otras muchas cosas con constancia y sin cansarse.

De vez en cuando aparecía algo ingenioso y distinto. Por ejemplo, con el telar Jacquard, que se inventó en 1801, se podían tejer complicadísimos tapices por medio de una sarta de tarjetas perforadas. En los albores del siglo XX, la electrónica había otorgado a las máquinas industriales unos sentidos limitados. Podían detenerse cuando algo no funcionaba bien y controlar la temperatura, el espesor e incluso la consistencia de las piezas que fabricaban. Sin embargo, cada máquina desempeñaba una tarea y sólo una. En consecuencia, el producto fabricado solía quedar obsoleto antes de que se hubieran amortizado los costes de diseño y construcción de la máquina. Este problema se hizo más grave después de la Segunda Guerra Mundial.

En 1954, el inventor George Devol patentó un nuevo tipo de máquina industrial, el brazo robot programable. Se podían modificar sus funciones sólo con alterar el orden de las tarjetas perforadas que controlaban sus movimientos. En 1958, junto con Joseph Engelberger, Devol fundó una empresa llamada Unimation (contracción de *universal* y *automation)* para construir estas máquinas. Las tarjetas perforadas pronto dieron paso a la memoria magnética, que permitía programar al robot llevándolo «de la mano» por el camino que tenía que recorrer, para expresarlo en términos corrientes. El primer robot industrial empezó a funcionar en 1961 en una de las fábricas de General Motors. Y hasta la fecha, los robots más grandes, los que hacen soldaduras, pintan con pistola y mueven las piezas de los coches, son de este tipo.

Pero los estudios sobre robótica que se hacían en las universidades empezaron a tener influencia sobre el diseño de los

robots industriales solamente cuando el precio de éstos bajó de los diez mil dólares. A finales de los setenta aparecieron los primeros sistemas industriales con visión, capaces de localizar e identificar piezas en las cintas transportadoras. Por lo general, llevaban acoplados unos pequeños brazos robot. Los robots que, en cierto modo, pueden ver y sentir juegan hoy un papel modesto —pero que va ganando en importancia— en los procesos de ensamblaje y de inspección de dispositivos pequeños, tales como calculadoras, cuadros de circuitos impresos y bombas de agua de automóviles. Una vez más, las necesidades industriales han ejercido una gran influencia sobre las investigaciones universitarias. Lo que en su momento fue un número insignificante de proyectos sobre robots, ha invadido los campus de las universidades de todo el país. Y aunque la cibernética pueda estar ahora relativamente inactiva, su robusta madre, la teoría de control, ha estado en actividad desde la guerra para satisfacer las lucrativas necesidades de la industria aeroespacial. Y en la actualidad existen elaborados métodos de control para aviones, astronaves y armas que dejan sentir su influencia sobre el diseño de los robots industriales.

En 1987 me invitaron a hacer una visita a la fábrica de Fremont, California, en la que se montan los ordenadores Apple Macintosh. Me di cuenta de que la mayor parte de la planta estaba bien organizada, pero que no era nada extraordinario. Muchas de las fases de montaje se hacían a mano. Posiblemente, las máquinas más eficaces eran las que insertaban componentes en cuadros de circuitos. Estos «rellenadores de cuadros» funcionaban como las máquinas de coser: sacaban los componentes, ensartados en cintas como las municiones de una ametralladora, y los «cosían» en los cuadros de los circuitos impresos a una velocidad vertiginosa, de varios componentes por segundo, mientras el cuadro se movía de izquierda a derecha, hacia delante y hacia atrás, para estar en cada momento en la posición adecuada. Estas máquinas son maravillas del control computarizado y su coste lo compensa el alto volumen de producción, pero me dejaron una vaga sensación de desaso-

siego. Sin embargo, en una pequeña hornacina vi un dispositivo, bastante distinto, que insertaba componentes que las máquinas de alta velocidad no podían manejar. Las piezas eran inductores antiguos, es decir, pequeños botes metálicos que contenían un carrete de alambre. Cada uno de los botes tenía unas lengüetas metálicas para encajar en las ranuras del tablero, y los carretes terminaban en alambres para los orificios pequeños. A diferencia de los componentes de las cintas de alimentación de las otras máquinas, hechos con toda precisión, los inductores, que se suministraban pulcramente ordenados en bandejas de plástico, tenían con frecuencia las lengüetas ligeramente dobladas y los alambres se podían estropear en cuanto fallara el intento de introducirlos en el tablero.

La máquina de inserción funcionaba en el interior de una cabina de cristal. Los tableros y las bandejas de inductores entraban y salían por medio de cintas transportadoras. El proceso de inserción comenzaba con una cámara de televisión enfocada sobre la bandeja de piezas. Un programa de visión localizaba un inductor y un rápido brazo robot lo cogía, lo levantaba y, con el extremo del alambre y la lengüeta hacia arriba, lo colocaba delante de otra cámara de televisión. Un segundo programa de visión examinaba las lengüetas y, si estaban rectas, dirigía el brazo para que insertara el componente en el tablero. Si las lengüetas estaban ligeramente curvadas, el inductor se llevaba primero a un «enderezador de pernos», un bloque metálico con unos orificios cónicos, y a continuación los pernos se examinaban de nuevo. Si los pernos estaban tan estropeados que no se podían enderezar, el inductor se tiraba a un cubo de desechos y se tomaba otro de la bandeja.

El proceso de inserción era bastante delicado. El inductor se colocaba rápidamente a unos pocos milímetros de la superficie del cuadro y se iba bajando lentamente hasta que el brazo robot encontraba resistencia. El brazo empujaba ligeramente el inductor hacia delante y hacia atrás y, al mismo tiempo, hacía una pequeña presión hacia abajo hasta que las lengüetas y los alambres encontraban los orificios que les correspondían. En-

tonces se aplicaba una presión mayor para ajustar el componente. Una recortadora, movida por un motor, cortaba y doblaba las protuberancias del metal, sujetando el inductor. Si después de algunos segundos la pieza no se podía ajustar, se colocaba delante de la cámara para que comprobara los cables y, si era necesario, se volvía a llevar al enderezador de pernos y se insertaba. Si el tercer intento fracasaba, entonces se tiraba la pieza al cubo de desechos y se cogía otra.

Todo esto tenía lugar con mucha rapidez. Cada tres o cuatro segundos se insertaba una pieza, aunque, si era defectuosa, podía necesitar hasta diez. Me sentí impresionado y un tanto nostálgico. Estas actividades me resultaban familiares. Una década antes, en el Laboratorio de Inteligencia Artificial de Stanford (SAIL), universidad en la que hice mis estudios de graduado, había sido testigo de procesos semejantes, algo más lentos y torpes. De hecho, el sistema de ensamblaje de Apple lo había creado una pequeña empresa del sur de California, llamada Adept, y tuvo sus orígenes en las tesis doctorales realizadas en el SAIL. Las semillas estaban empezando a germinar.

El objetivo de conseguir que los robots fijos actuaran de forma semejante a la de los seres humanos, aunque muy diluido por la miríada de objetivos a corto plazo y de modos de enfocar la cuestión, había adquirido un empuje implacable, darwiniano. Como historia, desconcierta tanto por su diversidad como por las interrelaciones que se producen. Entremos en el poco poblado mundo de los robots errantes.

MÁQUINAS QUE EXPLORAN (DE FORMA VACILANTE)

Los primeros programas de razonamiento necesitaban pocos datos para funcionar. Un tablero de ajedrez o un problema de lógica, geometría o álgebra, se podían describir mediante unos pocos cientos de símbolos bien elegidos. Y, de la misma manera, las reglas para resolver el problema se podían expresar con unos cientos de las llamadas «transformaciones» de estos

datos. La única dificultad residía en encontrar, de entre el número astronómico de posibles combinaciones, una secuencia de transformaciones que fuera la solución del problema. Pero era evidente que, para campos menos restringidos (como ejemplo retórico se solía usar la pregunta «¿Cómo se va a Tombuctú?»), se necesitaría un almacén de información mucho mayor. Y lo que no parecía probable es que se pudieran incluir manualmente en los programas todos los datos precisos para resolver estos problemas.

Los programas podrían conseguir algunos de esos datos por sí mismos, siempre y cuando aprendieran a leer libros y a entenderlos. Pero para comprender simples palabras es necesario tener un conocimiento detallado del mundo físico. Se supone que este conocimiento existe previamente en la mente del lector, ya que en ningún libro viene una descripción exhaustiva de lo que es una roca, un árbol, el cielo o un ser humano. Es posible que parte de este *conocimiento del mundo*, como se ha denominado, lo pudiera obtener la propia máquina si fuera capaz de observar directamente lo que le rodea por medio de ojos de cámara, oídos de micrófono y manos de robot dotadas de sensibilidad. Una de las primeras razones para elaborar proyectos de robótica en los laboratorios de inteligencia artificial fue el deseo de conseguir automáticamente este conocimiento del mundo. El modelo interno del mundo que elaboraran estos ordenadores podrían utilizarlo ellos mismos y otras máquinas, como fundamento para razonar con sentido común.

Aunque una máquina que se mueve pueda reunir muchos más datos que otra que esté inmóvil, la dificultad logística que suponía conectar un enorme ordenador inmóvil a un complicado conjunto de sensores situados sobre una plataforma móvil hizo que la solución de los sistemas de manos y ojos fijos les pareciera más conveniente a muchos investigadores. Además, pronto se descubrió que era mucho más complejo el problema de cómo adquirir informaciones del entorno que las propias actividades mentales para las que se querían utilizar estas informaciones. En los años setenta había docenas de laboratorios que tenían brazos robot conectados a ordenadores, pero casi

ninguno tenía robots móviles. La mayor parte de los investigadores del campo de la robótica consideraba que la movilidad era una complicación innecesaria que se añadía a un problema ya de por sí difícil de resolver. Su experiencia era totalmente opuesta a la de los cibernéticos (y a la de cientos de aficionados y fabricantes de juguetes), que se sentían bastante satisfechos si conseguían que los modestos circuitos de sus maquinitas móviles imitaran comportamientos muy parecidos a los de los animales.

En 1969 se terminó en el Instituto de Investigación de Stanford la construcción de Shakey, que fue el primer y único robot móvil hasta la fecha controlado fundamentalmente por medio de programas de razonamiento. Es la excepción que confirma la regla. Los que lo hicieron posible —Nils Nilsson, Charles Rosen y otros— se inspiraron en los primeros éxitos de las investigaciones sobre inteligencia artificial. Aplicaron los métodos de resolución de problemas lógicos a tareas del mundo real, en las que participaba una máquina capaz de moverse y de percibir su entorno. Para los diseñadores, los problemas que surgieron para controlar estos movimientos e interpretar las percepciones eran secundarios, ya que en lo que estaban principalmente interesados era en la capacidad de razonamiento de la máquina. De hecho, la tarea de producir el *software* que asegurara su movilidad y sensibilidad se confió a programadores de inferior categoría.

Shakey, que media un metro y medio y se movía impulsado por dos motores lentos pero precisos, llevaba una cámara de televisión. Otro ordenador, muy grande, lo atendía por control remoto. Para interpretar las imágenes de televisión se adaptaron los métodos del MIT de manipulación de programas por bloques, que con anterioridad sólo se habían usado con brazos robot. Únicamente funcionaban cuando los objetos captados eran simples, planos y de un solo color, así que se construyó para el robot un entorno especial que constaba de varias habitaciones con paredes despejadas, en las que había un cierto número de bloques y cuñas pintados de un solo color. La hazaña más impresionante de Shakey, realizada por etapas a lo largo

de varios días, fue resolver el problema «del mono y los plátanos». Se le ordenó que empujara un cierto bloque que estaba colocado sobre otro mayor, y el robot elaboró y siguió un plan según el cual tenía que encontrar una cuña que le sirviera de rampa, empujarla hasta que estuviera pegada al bloque más grande, subir la rampa y empujar el bloque pequeño tal como le habían mandado.

Tanto el problema como el entorno eran inventados, pero proporcionaron una motivación y la comprobación de un programa inteligente de razonamiento llamado STRIPS (Stanford Research Institute Problem Solver, es decir, el Solucionador de Problemas del Instituto de Investigación de Stanford). Se le pedía al robot que realizara una tarea, y STRIPS elaboraba un plan basado en los limitados movimientos que podía efectuar el robot, cada uno de los cuales tenía condiciones previas (por ejemplo, para empujar un bloque, lo tengo que tener delante de mí) y consecuencias probables (por ejemplo, después de empujar un bloque, se mueve). El estado del mundo del robot se representaba por medio de frases de lógica matemática, y elaborar un plan era lo mismo que demostrar un teorema: el estado inicial del mundo eran los axiomas, los primeros movimientos eran las reglas de inferencia, y el resultado deseado era el propio teorema. Pero inmediatamente apareció una complicación: el resultado de una acción no siempre es el que uno espera. Por ejemplo, cuando el bloque no se mueve. La capacidad de Shakey para enfrentarse a estas irregularidades era limitada, ya que de cuando en cuando observaba partes del mundo y tenía que ajustar sus descripciones internas y volver a plantearse sus movimientos si esas condiciones no eran las que había supuesto.

Shakey era impresionante como concepto, pero digno de compasión cuando se movía. Cada uno de los movimientos del robot, cada ojeada que daba con su cámara, consumía aproximadamente una hora de tiempo de ordenador y entrañaba una alta probabilidad de fracaso. El ejercicio de empujar los bloques, que se ha descrito más arriba, se llevó a cabo y se filmó por etapas, y se necesitaron varias «tomas» para cada una de

las etapas antes de que quedara bien. Los planificadores de STRIPS no tuvieron la culpa del fracaso, ya que elaboraban buenos planes cuando se les proporcionaban buenas descripciones del entorno del robot. La culpa fue de los programas que interpretaban los datos sin procesar que les proporcionaban los sensores y actuaban de acuerdo con las recomendaciones.

En los primeros años de la década de los setenta me parecía que algunos de los creadores de los programas de razonamiento que funcionaban bien sospechaban que los malos resultados de los trabajos de robótica reflejaban, de alguna forma, la capacidad intelectual de los que estaban en ese campo de la investigación. Este esnobismo intelectual no es raro, por ejemplo, entre los físicos teóricos y experimentales. Pero a medida que ha ido creciendo el número de demostraciones, se ha visto que es relativamente sencillo hacer que los ordenadores tengan la capacidad de los adultos para resolver problemas, *tests* de inteligencia o jugar a las damas, pero que es difícil, por no decir imposible, dotarles de la capacidad de percepción y movilidad de un niño de un año.

Si miramos atrás, esta dicotomía no resulta sorprendente. Desde que aparecieron los primeros animales multicelulares, hace unos mil millones de años, el animal capaz de realizar el movimiento adecuado después de una percepción confusa, es el que ha sobrevivido a la feroz lucha por recursos tan limitados como el espacio, la comida o la pareja. En las zonas sensoriales y motoras del cerebro humano, grandes y muy evolucionadas, está codificada la experiencia acumulada durante mil millones de años sobre la naturaleza del mundo y la manera de sobrevivir en él. En mi opinión, el proceso deliberado que denominamos razonamiento es el barniz más superficial del pensamiento humano, y es efectivo solamente porque recibe el apoyo de este conocimiento sensomotor, mucho más antiguo y mucho más fuerte aunque, por lo general, inconsciente. Todos nosotros somos prodigiosos campeones olímpicos en las áreas de la percepción y del movimiento, hasta tal punto que hacemos que lo difícil parezca fácil. Sin embargo, el pensamiento abstracto es

un truco nuevo, ya que debe tener menos de cien mil años. Todavía no lo dominamos. Y no es porque sea intrínsecamente difícil, sino porque nosotros hacemos que lo parezca.

Parece ser que los organismos que no tienen capacidad para percibir y explorar su entorno carecen de lo que denominaríamos inteligencia. Simplemente comparando el reino animal y el vegetal podemos observar que los organismos que se mueven tienden a desarrollar las características mentales que asociamos con la inteligencia, y que los que no se mueven no las tienen. Las plantas son dignas de respeto por su efectividad, pero no parecen mostrar ninguna inclinación hacia la inteligencia. A lo mejor dentro de muchísimo tiempo existirán plantas inteligentes —algunas plantas carnívoras y «sensibles» son la prueba de que es posible que se dé algo parecido a un acto nervioso—, pero las expectativas de vida del universo no son tan amplias.

Los investigadores del campo de la cibernética, cuyos experimentos independientes con frecuencia se inspiraban en la movilidad animal, empezaron a investigar los sistemas nerviosos para intentar duplicar las capacidades sensoriales y motoras de los animales. El grupo que se dedicaba a la inteligencia artificial ignoró este enfoque en sus primeros trabajos. Por el contrario, puso sus aspiraciones en la cumbre de la perfección, el pensamiento humano, y se dedicó a hacer experimentos con grandes ordenadores de alta velocidad estacionaria para mecanizar el razonamiento puro. Este itinerario «de arriba abajo» hacia la inteligencia de las máquinas permitió que se hicieran impresionantes avances en un principio, pero, en conjunto, los logros conseguidos en una década son poquísimos. Mientras la cibernética arañaba la superficie inferior de la inteligencia real, la inteligencia artificial arañaba su parte superior. Y el núcleo interior permanecía sin tocar.

Todos los intentos para conseguir que las máquinas sean inteligentes han pasado por la fase de imitación de la inteligencia natural, pero los distintos modos de enfocar el problema han remedado diferentes aspectos del original. Lo que pretende

la inteligencia artificial tradicional es copiar los procesos mentales conscientes de los seres humanos cuando llevan a cabo distintas tareas. Pero esto tiene una limitación, y es que los aspectos más poderosos del pensamiento son inconscientes, inaccesibles a la introspección mental y, en consecuencia, es difícil registrarlos formalmente. Algunos cibernéticos tomaron otro rumbo y se dedicaron a construir modelos de sistemas nerviosos de animales desde el nivel neurológico. Pero esto es muy limitado, ya que los sistemas nerviosos grandes tienen un número astronómico de células y es muy difícil determinar con exactitud qué es lo que hacen las neuronas individuales, cómo están interconectadas y cómo funcionan las redes nerviosas. Tanto la IA tradicional como los modelos neurológicos han aportado muchas ideas y, con tiempo suficiente, cualquiera de las dos podría resolver el problema. Pero, tal y como están las cosas, creo que la manera más rápida de avanzar es imitar la *evolución* de las mentes de los animales y, poco a poco, ir añadiendo capacidades a las máquinas, de forma que la secuencia resultante de comportamientos de la máquina se asemeje a las aptitudes de animales con sistemas nerviosos cada vez más complicados. Una de las características fundamentales de este enfoque es que se puede ajustar la complejidad de estos avances paulatinos para aprovechar mejor la capacidad de resolver problemas de los investigadores y de los ordenadores. Nuestra inteligencia, como herramienta, nos debe permitir seguir la trayectoria de la inteligencia a un ritmo más rápido que el de los pacientes pero ciegos procesos de la evolución darwiniana.

El camino a seguir es de abajo arriba, y los primeros problemas son los de la percepción y la movilidad, porque la inteligencia humana se desarrolló apoyándose en la sólida roca sensomotora. Se están creando programas que plantean problemas similares a los que tuvieron que resolver los animales primitivos: cómo afrontar e incluso anticipar las sorpresas, peligros y oportunidades que se le presentan a un organismo que explora su entorno. Pero se comprueban en robots que tienen que enfrentarse con las incertidumbres de un mundo real. La mayor parte de estos programas fracasarán, pero unos pocos

tendrán éxito, lo mismo que sucede con la diminuta fracción de mutaciones espontáneas de los organismos que sobrevive a la siguiente generación. La supervivencia depende de las posibilidades que ofrezca una técnica nueva para hacer frente al desafío de un entorno complejo y dinámico. Si planteamos condiciones experimentales análogas a las que se encontraron los animales en el curso de su evolución, es posible que podamos seguir los pasos de la evolución de la inteligencia humana. El hecho de que los animales comenzaran con sistemas nerviosos reducidos nos infunde confianza. Es probable que los pequeños ordenadores de hoy puedan emular los primeros pasos hacia comportamientos semejantes a los de los seres humanos. Siempre que sea posible, daremos una ojeada biológica a la «parte final del libro», y nos servirá de ayuda en nuestro empeño de simular la inteligencia. La «parte final del libro» son las características de los animales y los seres humanos en el campo neurológico, morfológico y de comportamiento, que han sistematizado los que estudian estos aspectos de la vida.

Los trabajos en el campo de la moderna robótica cuentan apenas veinte años, y solamente durante los diez últimos se han utilizado ordenadores para realizar el control rutinario de los robots. El resumen de la evolución de la vida inteligente está todavía en una fase muy temprana. Existen equivalentes robóticos de sistemas nerviosos, pero, por lo 'que se refiere a su complejidad, son comparables a los de los gusanos. Sin embargo, las presiones evolutivas que determinaron la vida ya se producen en los laboratorios de robótica, y creo que este itinerario hacia la inteligencia artificial, de abajo arriba, un día se encontrará con el tradicional, el que va de arriba abajo. Es posible que este encuentro se produzca algo más allá de la mitad del camino, y que proporcione la eficacia y el conocimiento lógico que tan esquivo se ha mostrado con los programas de razonamiento. Y habrá máquinas totalmente inteligentes cuando la grapa de oro metafórica una estas dos trayectorias. Un programa de razonamiento al que respalde un modelo robótico del mundo podrá visualizar los pasos que debe dar en su plano, distinguir las situaciones razonables de las absurdas e

intuir soluciones a partir de la observación, lo mismo que hacen los seres humanos. Después explicaré por qué espero que esta unión se produzca en los próximos cuarenta años. Por el momento, vamos a explorar parte de la fauna artificial que se encuentra en la parte de abajo.

Como ya hemos visto, Shakey no fue simplemente uno más. Este robot fue un ejemplo de la trayectoria de arriba a abajo. Su especialidad era razonar, mientras que el *software* rudimentario que se empleó para su visión y su movimiento únicamente funcionaba en entornos muy simples. Más o menos por la misma época, y con un presupuesto mucho más bajo, como parte del Proyecto de Inteligencia Artificial de la Universidad de Stanford, a unos 13 kilómetros de la residencia de Shakey, nació un robot móvil que se iba a especializar en ver y moverse en escenarios naturales. John McCarthy, que fundó el Proyecto en 1963, perseguía el objetivo, entonces plausible, de construir en diez años una máquina totalmente inteligente. Cuando terminó la década y, con ella, la plausibilidad del objetivo, se le dio un nuevo nombre: Laboratorio de Inteligencia Artificial de Stanford o SAIL. McCarthy, haciéndose eco de las prioridades de las primeras investigaciones que se realizaron en el campo de la inteligencia artificial, trabajó sobre el razonamiento y delegó en otros el diseño de los oídos, los ojos y las manos de esta mente artificial. El grupo de SAIL dedicado a los ojos y manos en seguida adelantó al grupo de robótica del MIT, y fue el germen del auge que experimentó posteriormente la fabricación de brazos de robot, que se dedicaron a la industria. Se añadió una pequeña ayuda a la movilidad cuando Les Earnest, el astuto administrador jefe de SAIL, descubrió un vehículo que había abandonado el departamento de ingeniería mecánica y que, tras unas ligeras modificaciones, funcionaba como un explorador lunar accionado por control remoto. El SAIL se convirtió en el Carro de Stanford, el primer robot móvil controlado por un ordenador grande que *no* razonaba, y el primer banco de pruebas para la visión por ordenador en el mundo confuso y mal iluminado que habita la mayoría de los

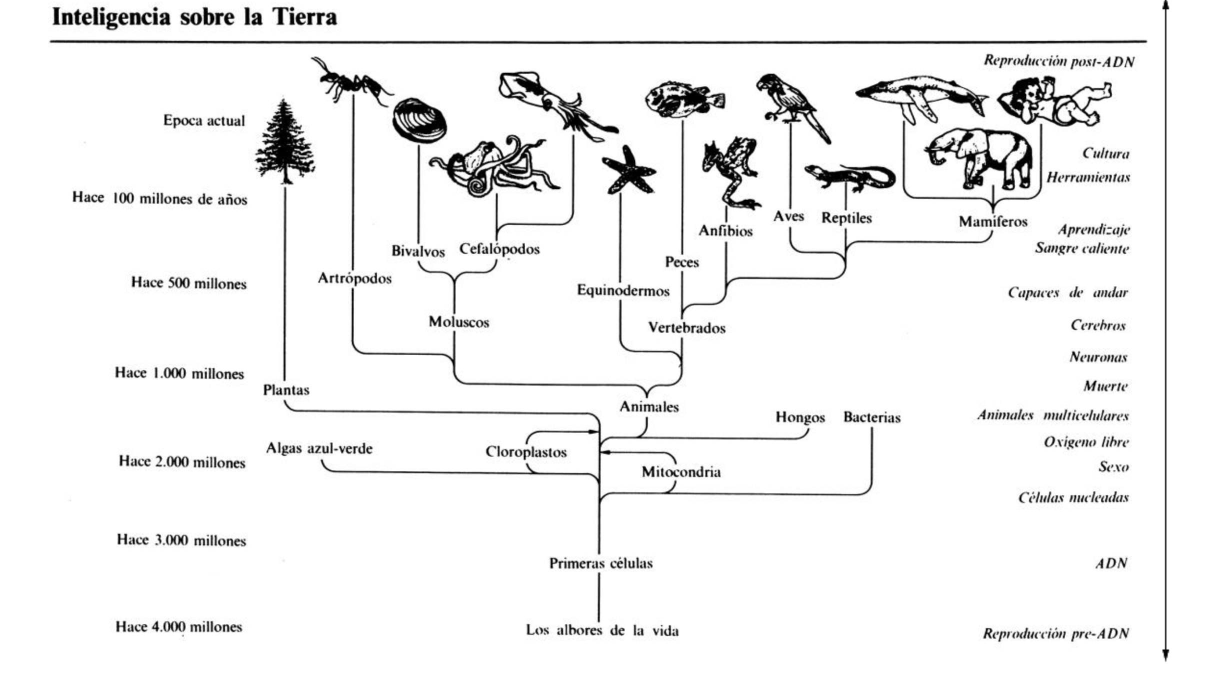
animales. El Carro de Stanford, descendiente de dos tesis doctorales (una de ellas, la mía), se movía por espacios exteriores e interiores guiado por imágenes de televisión que procesaban programas muy distintos de los del mundo de Shakey.

A mediados de la década de los setenta, la NASA empezó a proyectar una misión robot a Marte, como consecuencia de los éxitos de los aterrizajes del Viking. El lanzamiento se programó para 1984, y la misión estaría compuesta por dos vehículos que explorarían la superficie marciana. Este planeta está tan lejos, incluso para la radio, que el control remoto simple resultaría demasiado lento o arriesgado. El lapso de tiempo transcurrido entre el envío de una orden y el momento de poder ver sus resultados puede ser de cuarenta minutos. Si el robot pudiera moverse por sí mismo, con seguridad, la mayor parte del tiempo, podría cubrir mucho más terreno. El Laboratorio de Propulsión a Reacción de Caltech (JPL), que había diseñado la mayor parte de las astronaves robots de la NASA y que hasta entonces había utilizado un sistema de automatización bastante seguro y sencillo, comenzó a elaborar un proyecto robótico inteligente. Reunió métodos, *hardware* y gente que hubiera trabajado en programas robóticos universitarios, y se construyó una gran plataforma de prueba, con ruedas, que recibió el nombre de Vehículo de Investigación Robótica (RRV). El RRV llevaba cámaras, un telémetro de láser, un brazo de robot y un equipo electrónico completo, todo ello conectado a un ordenador grande por medio de un largo cable. En 1977 podía avanzar por tramos cortos en un aparcamiento lleno de piedras, coger una en concreto y girarla ante las cámaras. Pero en 1978 el proyecto se interrumpió porque se canceló la misión a Marte de 1984 y se eliminó del presupuesto de la NASA. Pero Marte no ha desaparecido, y JPL está pensando en ir de visita allí a finales del milenio.

El primer grupo partidario, y el más firme, de la investigación sobre la inteligencia artificial, junto con el Departamento de Investigación Naval, pertenece al Departamento de Defensa: la Agencia para Proyectos Avanzados de Investigación (DARPA). Entre otras razones, porque los primeros avances en

este campo se produjeron en los Estados Unidos. Se fundó en 1958, después de la humillación nacional que supuso el Sputnik, y su objetivo era consolidar proyectos a largo plazo como un seguro contra sorpresas tecnológicas desagradables. En 1981, los directivos del DARPA decidieron que la navegación robot ya había avanzado lo suficiente como para que mereciera la pena hacer un intenso esfuerzo y crear vehículos autónomos que pudieran recorrer largas distancias por tierra sin un operador humano, acaso por zonas de guerra o conflictivas. El número de proyectos para robots móviles creció vertiginosamente en las universidades y entre los proveedores de la defensa, que proporcionaron fondos para que el proyecto se hiciera realidad. Incluso ahora, hay algunos robots del tamaño de un camión que todavía están en período de prueba por todo el país... y las aguas no han vuelto aún a su cauce.

A otro nivel más cotidiano, no es ninguna insignificancia que, en las fábricas, se asignen ciertas tareas a brazos de robot fijos. Una solución sería una línea de montaje con cinta transportadora, pero los directores de las fábricas cada vez más automatizadas de finales de los setenta y principios de los ochenta pensaban que estas cintas limitan mucho, ya que no es nada sencillo cambiar sus trayectorias. Los robots se podían programar de nuevo para que desempeñaran distintas tareas, pero no se podía hacer lo mismo con los caminos que seguía el material. Varias multinacionales resolvieron el problema construyendo lo que denominaron Vehículos Guiados Automáticamente (GAV). Estos GAV se movían cuando recibían las señales que se les hacían llegar por medio de cables enterrados a lo largo de su camino. Parecían elevadoras o coches de choque grandes, se podían programar para que se movieran de un sitio a otro, y las operaciones de carga y descarga las realizaban brazos robot. Pero enterrar los cables en el suelo de cemento de las fábricas resulta caro, así que se buscaron otros métodos alternativos para que se movieran. Lo mismo que en el caso de los brazos robot, para crear robots móviles se han unido los esfuerzos académicos e industriales y se está llevando a la práctica un número pasmoso de ideas.


Inteligencia sobre la Tierra
Epoca actual
Hace 100 millones de años
Hace 500 millones
Hace 1.000 millones
Hace 2.000 millones
Hace 3.000 millones
Hace 4.000 millones
Bivalvos
Cefalópodos
Artrópodos
Moluscos
Equinodermos
Peces
Anfibios
Aves
Reptiles
Mamíferos
Vertebrados
Plantas
Animales
Hongos
Bacterias
Algas azul-verde
Cloroplastos
Mitocondria
Primeras células
Los albores de la vida
Reproducción post-ADN
Cultura
Herramientas
Aprendizaje
Sangre caliente
Capaces de andar
Cerebros
Neuronas
Muerte
Animales multicelulares
Oxígeno libre
Sexo
Células nucleadas
ADN
Reproducción pre-ADN

Este árbol genealógico parcial de los organismos terrestres nos sugiere la relación que existe entre movilidad e inteligencia. Hace mil quinientos millones de años, nuestros antepasados unicelulares estaban en compañía genética de las plantas. Como ambos eran unicelulares, nadaban libremente, pero cuando las plantas pasaron a ser multicelulares, se hicieron sedentarias y se limitaron a aprovechar la energía solar. Por otro lado, nuestros ancestros animales siguieron siendo ambulantes y se alimentaban de las plantas y de sus semejantes. Mientras las plantas prosperaban muchísimo —la vegetación es la mayor parte de la biosfera de la Tierra y las plantas son los organismos más numerosos y que viven más tiempo—, los animales no parecían tener ninguna tendencia a evolucionar hacia nada que pudiéramos denominar inteligencia.

Es cierto que los artrópodos —insectos, arácnidos y otros seres con muchas patas que avanzan arrastrándose— se mueven. Tienen órganos sensoriales muy desarrollados (lo mismo se puede decir de su sistema nervioso) y exhiben una complejidad de comportamientos parecida a la de los robots de la actualidad. Pero debido a su exoesqueleto, sufren limitaciones por lo que se refiere a su metabolismo y a su tamaño, y sus sistemas nerviosos nunca tienen más de un millón de neuronas, que es una cifra muy baja si se compara con los diez millones que tienen los cerebros de los vertebrados más pequeños. Los insectos sociales han logrado superar parcialmente las limitaciones que impone el tamaño coordinándose en amplios grupos de individuos que actúan como si fueran un solo animal.

Los moluscos están tan lejos de nosotros como los artrópodos: nuestros antepasados comunes existieron hace unos mil millones de años. Tienen un interés especial debido al contraste de estilos de vida entre las distintas especies. La mayor parte de los moluscos son crustáceos que se mueven lentamente e ingieren el alimento por medio de filtros, pero los cefalópodos —pulpo, calamar, jibia y nautilo— nadan libremente. El pulpo y el calamar perdieron la concha y son los invertebrados que tienen más alto grado de movilidad y sistemas nerviosos más grandes. Lo mismo que los vertebrados, poseen ojos capaces de formar imágenes, pero esos ojos son semiesferas unidas al tejido corporal flexible y las fotocélulas de la retina están dirigidas hacia el cristalino en vez de hacia afuera, como les sucede a los invertebrados. El cerebro tiene ocho ganglios colocados en forma de anillo alrededor del esófago. Es impresionante ver nadar tanto al pulpo como a sus parientes. Su cuerpo está recubierto por un millón de células que cam-

bian de color y se pueden controlar individualmente. Los pulpos pequeños son capaces de aprender a resolver algunos problemas como, por ejemplo, abrir un recipiente que contenga alimento. El calamar gigante es del tamaño de la ballena y su sistema nervioso es muy grande. Se conoce su existencia porque se han encontrado flotando en el mar algunos cadáveres hinchados, pero vive permanentemente a profundidades tan grandes que imposibilitan la observación humana.

Los equinodermos, entre los que se cuentan la estrella de mar, el erizo de mar y otros habitantes del fondo marino con el cuerpo dividido en cinco partes simétricas, tienen una forma de vida estable y cazan animales sedentarios, como moluscos y corales. Aunque están íntimamente relacionados con los vertebrados, sus movimientos son mucho más lentos, sus órganos sensoriales mucho más simples y su sistema nervioso mucho más pequeño. Cuando una estrella de mar caza una almeja, a los ojos de los seres humanos parece que está dormida, pero en una secuencia de fotos se puede ver que acecha a su presa, la ataca súbitamente, la abre y empieza a digerirla después de llevar su estómago hasta la concha abierta.

Los peces son muy móviles y sus cerebros suelen ser mayores que los de casi todos los invertebrados, pero menores que los de los otros vertebrados, posiblemente a causa de que el agua es un hábitat simple. Los reptiles evolucionaron sobre la tierra y su sistema nervioso es mayor. Los anfibios son animales intermedios entre los peces y los reptiles, por lo que se refiere al tamaño del sistema nervioso y a otras características. Los animales de sangre fría son de metabolismo lento y no pueden mantener sistemas nerviosos tan grandes como los de los pájaros o los mamíferos, pero son capaces de sobrevivir con un estilo de vida menos frenético.

Las aves están relacionadas con nosotros a través de un primitivo reptil que vivió hace unos trescientos millones de años. Las necesidades del vuelo limitan su tamaño, pero tienen un metabolismo muy activo. Aunque su estructura es distinta de la de los mamíferos, tienen el cerebro tan grande como el de mamíferos del mismo tamaño que ellas. También son semejantes por lo que se refiere a capacidad de aprendizaje y de resolver problemas.

Nuestro último antepasado en común con las ballenas fue un mamífero primitivo, parecido a la rata, que vivió hace cien millones de años. Algunas especies de delfines tienen masa corporal y cerebral idéntica a la nuestra, tienen la misma capacidad de resolver ciertos tipos de problemas que nosotros y pueden entender y comunicar ideas complejas. Los cachalotes tienen el cerebro más grande del mundo.

La ventaja que pueden tener los seres humanos sobre otros animales de cerebro grande, como elefantes y ballenas, no dependerá tanto de nuestra inteligencia individual como de la rapidez con que esa inteligencia se adapte a nuestra cultura, que utiliza herramientas, es inmensamente poderosa y evoluciona a gran velocidad.

UN ROBOT PARA LAS MASAS

Todo lo que se ha contado hasta ahora es bastante aséptico, y sólo se ha hablado de algunos de los actores principales del nuevo campo de la robótica. La realidad es como las pociones de las brujas: una mezcla de enfoques, motivaciones y problemas que, hasta la fecha, no se han podido relacionar. Los protagonistas son grupos, grandes y pequeños, de ingenieros (electrónicos, mecánicos, ópticos y de todo tipo), físicos, matemáticos, biólogos, químicos, especialistas en tecnología médica, científicos dedicados a los ordenadores, artistas e inventores de todo el mundo. Los especialistas en ordenadores y los biólogos colaboran en el campo de las máquinas que ven. Los físicos y los matemáticos trabajan para mejorar el sonar y otros sensores. Los ingenieros mecánicos han construido máquinas que caminan sobre piernas, y otras que son capaces de agarrar y que poseen unas manos de robot casi tan diestras como las humanas. Y, sin embargo, apenas ha existido la comunicación entre los distintos grupos. Han sido incapaces de llegar a un acuerdo y trazar un bosquejo general del campo de la robótica. A pesar del caos, espero que la producción en serie llegue a tiempo para que vea la entrada del milenio un robot multiuso, que se pueda utilizar tanto en la fábrica como en el hogar.

En las naciones industrializadas, cada vez se emplean más máquinas en la agricultura y en la industria, y la gente tiene más tiempo para dedicarse a realizar otras actividades humanas. Los alimentos y los bienes de consumo en general son más abundantes y más baratos, pero los costes de muchos servicios se han incrementado considerablemente. Por ejemplo, el servi-

cio doméstico es escaso y caro. Y las máquinas domésticas tales como los procesadores de alimentos, aspiradoras y hornos de microondas no llenan el vacío, especialmente en las familias en las que todos los adultos trabajan fuera del hogar. Esta necesidad se planteó hace varias décadas: *¿Cuándo existirá un robot para ayudar en las tareas domésticas?*

Durante muchos años he creído que era bastante improbable que existieran robots sirvientes, siempre presentes en la ciencia ficción sobre el futuro próximo. Las casas son entornos complejos y con recursos limitados. El rendimiento económico de un ayudante doméstico mecánico supondría solamente una pequeña parte del valor de un robot en un puesto industrial típico y, por lo tanto, los robots para el hogar se tendrían que vender (o que alquilar) por mucho menos. Pero no es eso todo. Sería mucho más complicado que funcionaran de una forma efectiva y segura en un entorno con frecuencia caótico, como una casa, que en una fábrica. Los robots actuales pueden hacer movimientos ciegos y repetitivos a un precio comparable al de toda la casa. Esta enorme diferencia entre el precio y el rendimiento es real y, sin embargo, espero que en el plazo de diez años exista un robot multiuso para el hogar. Este cambio de actitud es consecuencia, en parte, de los progresos de la investigación en los últimos años, y en parte también a que ahora entiendo de otra manera las implicaciones del concepto «multiuso».

Los robots industriales que existen hoy en día son mucho más flexibles que la automatización fija, a la que con frecuencia desplazan, pero no hacen bien muchas de las cosas a las que se puede aplicar el término «multiuso». Además, los robots individuales se suelen asignar a un puesto fijo, llevan agarradores y, en ocasiones, sensores para realizar una tarea específica, cosa que hacen una y otra vez acaso para el resto de sus vidas. Su repertorio es muy reducido y aburrido y, por lo tanto, el número de ellos que se puede vender es pequeño. Hoy en día existen en el mundo menos de cien mil robots, de todas las marcas, que no sean juguetes. Comparemos esta cifra con las siguientes: cien millones de coches, quinientos millones de aparatos

de televisión o doscientos millones de ordenadores. Como se venden tan pocas unidades de robots, no se les dedica mucho esfuerzo. Y el resultado es un diseño menos que óptimo a un precio muy elevado. Pero no será siempre así. A medida que crece el número de unidades que se producen, aumentan la oportunidad y los incentivos para mejorar el diseño de los robots y los detalles de su producción. Los costos son menores y los robots mejores, con lo que crece el mercado y, en consecuencia, el número de unidades vendidas. Y se producen más mejoras. La gráfica de la disminución de precio por unidad en relación al número de unidades producidas recibe el nombre de *curva de aprendizaje del fabricante*.

El mercado potencial de los robots crecerá enormemente cuando se alcance un cierto nivel de utilidad general. Hasta llegar a ese punto en que los beneficios igualen a las pérdidas, la especialización —la explotación de circunstancias únicas para conseguir un rendimiento aceptable con un mínimo de complejidad— será la norma de la robótica. Cuando se sobrepase este punto, el mercado potencial será bastante amplio y los beneficios más elevados tendrán como consecuencia diseños cada vez más estándar, que se venderán en grandes cantidades. Habrá llegado el robot barato, producido en serie y de gran utilidad. Hemos acumulado la suficiente experiencia como para poder detallar algunas de las características de este Modelo T de robots. No serán inteligentes ni vendrán programados previamente para poder realizar una gran cantidad de tareas útiles. Vendrán de la fábrica con un conjunto de habilidades mecánicas, sensoriales y de control que se podrán invocar a conveniencia mediante un *software* especialmente elaborado para aplicaciones específicas.

El primer mercado importante para estas máquinas serán las fábricas, donde resultarán algo más baratos y considerablemente más versátiles que la antigua generación de robots a la que sustituirán. Como la relación costo/beneficios será más satisfactoria, se podrán dedicar a muchas más tareas, con lo que se necesitarán más unidades y los costos descenderán. Con el

tiempo, llegarán a ser más baratos que un coche pequeño, estarán al alcance de bastantes hogares y crecerá la demanda para que se elabore un nuevo *software* mucho más amplio. Los programas de control de robots con distintas tareas vendrán de múltiples fuentes, lo mismo que sucede hoy en día con los programas para los ordenadores personales y comerciales.

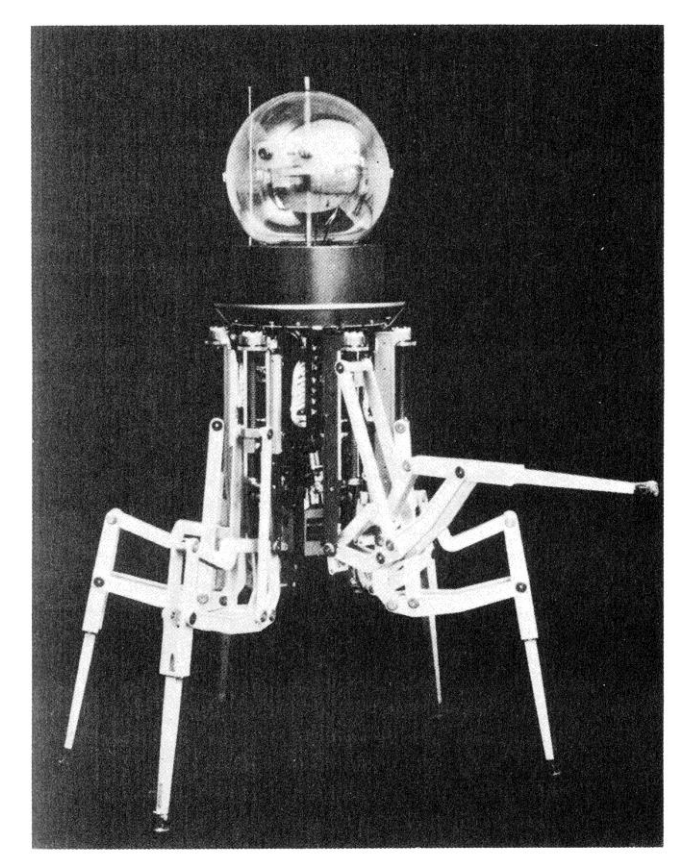
Y también como en el caso de los ordenadores personales, se encontrarán muchas aplicaciones para el robot multiuso que sorprenderán a los fabricantes. Podemos especular sobre los equivalentes a los videojuegos los procesadores de textos y los programas de contabilidad que existirán en la era de los robots fabricados en serie, pero la realidad es desconocida. Para seguir con el juego de las adivinanzas, pensemos en los programas que aligeran el montaje mecánico (de la fábrica de una empresa de automatización), limpian los cuartos de baño (de una pequeña empresa creada por antiguo personal de limpieza), hacen comidas propias de un gastrónomo con ingredientes frescos (colaboración entre un tipo de ordenador y un *chef* de París), pondrán a punto los coches Saturno de un cierto año (del departamento de servicios de Saturno de la General Motors), colgarán alfombras (de un estudiante de enseñanza media de Massachusetts), podarán el césped arrancando las malas hierbas una por una, participarán en carreras de robots (contra otros *softwares,* y los programas los asignará un robot físico por sorteo antes de que empiece la carrera), harán trabajos de excavación (de una empresa de construcción innovadora), investigarán amenazas de bomba (vendido a los departamentos de policía de todo el mundo), realizarán inventarios de almacenes, ayudarán a montar y a comprobar otros robots (en varias fases independientes), y muchas otras tareas. Para algunas de las aplicaciones, el robot precisará accesorios opcionales de *hardware*, herramientas especiales, sensores (como olfateadores químicos), cubiertas protectoras y otros.

Puede suceder que la actividad de escribir programas de aplicaciones para sucesivas generaciones de robots multiuso llegue a ser una de las principales ocupaciones del ser humano en las primeras décadas del siglo que viene. Por ejemplo, un

fontanero experimentado tendrá que elegir entre aplicar directamente su técnica para satisfacer las necesidades de varios cientos de clientes, o codificarla en programas de robot que podrá vender a miles o incluso a millones de compradores. La primera alternativa irá perdiendo paulatinamente sus encantos como fuente de ingresos, ya que el trabajo manual entrará en competencia con un número creciente de sofisticados robots controlados por un *software* cada vez más perfecto. Esto último también entraña sus riesgos. Un programa puede ser un rotundo fracaso, lo mismo que los inventos, la música, el arte y el *software* de los ordenadores en el mundo actual. Pero un programa puede también representar para su creador una fuente de ingresos durante muchos años.

Casi todo el mundo tiene o puede desarrollar muchas habilidades, y cada una de ellas es una fuente potencial de ingresos si se codifica como programa. Y se pueden sacar al mercado y poner a la venta muchas versiones competitivas de cada una de estas técnicas especiales basándose en la utilidad, el costo, el gusto personal, la moda y la publicidad. Los programas tienen una vida limitada y están condenados a que los eclipsen otros mejores o diseñados para una generación de robots más sofisticada. Y surgirá una importante industria secundaria para colaborar en los procesos de programación.

Antes de que transcurra mucho tiempo se crearán programas que convertirán a los robots multiuso en buenos aprendices, y se les podrá enseñar a llevar a cabo una cierta tarea mostrándoles bien los movimientos que tienen que hacer o un ejemplo de los mismos. La biblioteca en la que se reúnan estos programas será, finalmente, como un almacén en el que se encontrará codificado todo el conocimiento humano no verbal, que podrán aprovechar las nuevas generaciones de robots autónomos que sigan a la primera generación viable. La industria de sistemas expertos ya ha empezado a codificar de esta manera el conocimiento verbal.



La máquina que anda
La Odex de Odetics puede caminar, saltar, estrecharse para atravesar puertas o expandirse para conseguir más estabilidad. Pero su alto consumo de energía hace que haya que cambiar de baterías sólo después de una hora de movilidad.

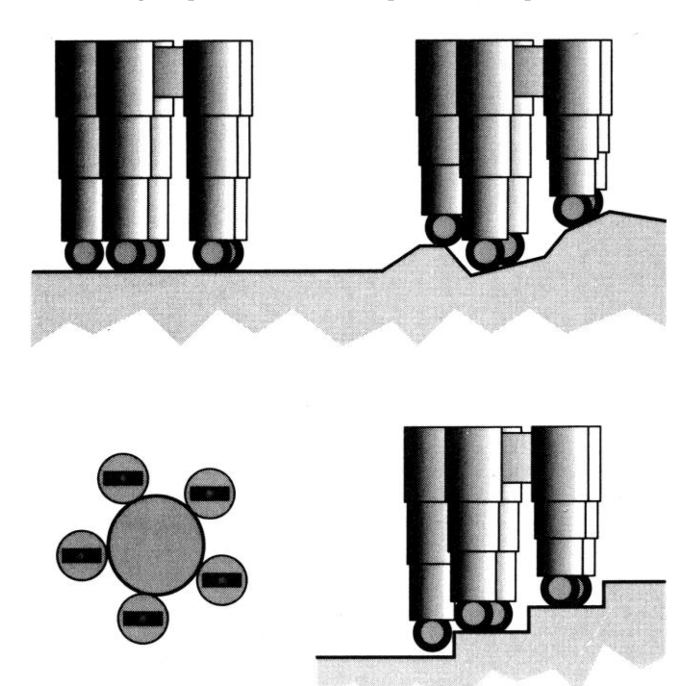
LOCOMOCIÓN EFICIENTE

Para que resulte rentable, un robot multiuso producido en serie deberá tener un nivel mínimo de funcionalidad. Los primeros robots de este tipo no tendrán que saber hacer todo bien, ni siquiera muchas cosas. Deberán hacer bastantes cosas suficientemente bien como para que se puedan abrir un mercado no limitado de antemano, de forma que cada vez que bajen los precios se producirá un incremento proporcional en la demanda del número de aplicaciones económicas y de unidades.

Incluso en las zonas totalmente urbanas, las superficies de suelo plano y duro forman un archipiélago en un mar de terreno que es áspero o suave, que se puede hollar o es imposible de atravesar. Una máquina que no pueda navegar por este mar se quedará bloqueada en una isla y sus usos potenciales se verán enormemente restringidos. Nuestro criterio de eficiencia exige, por lo tanto, que el sistema del mecanismo de trasmisión sea mejor que las ruedas corrientes. Y parece que los robots con piernas están demostrando ser los más prácticos. La demostración más convincente, hasta la fecha, la ha proporcionado Odetics, una empresa de California: ha creado un robot con seis piernas, como una araña, que se mueve por medio de un sistema eléctrico de tracción, y que puede salir del vehículo que lo transporta, subir a una sombrerera, atravesar una puerta estrecha y despertar la admiración de todos por su capacidad de levantar uno de los extremos del vehículo que lo transporta y arrastrarlo. Ésta y otras pruebas nos hacen pensar que, probablemente dentro de una década, sea asequible la locomoción práctica con piernas.

Las piernas son un mecanismo de movimiento excelente, pero su propia naturaleza limita la velocidad y el rendimiento de la energía. Por ejemplo, la máquina de Odetics agota las baterías en menos de una hora si se mueve despacio. Las serias limitaciones de potencia en un robot independiente van a exigir que el mecanismo de trasmisión sea más frugal. En suelo plano, lo mejor son las ruedas, ya que ofrecen cerca del cien por cien

de eficacia para una amplia escala de velocidades. Una solución intermedia serían unas piernas terminadas en pies con ruedas, parecidos a unos patines con motor. El robot podría moverse con estos pies la mayor parte del tiempo y levantarlos cuando tuviera que salvar obstáculos o subir escaleras. En suelo desigual podría avanzar despacio dando pasos cortos.



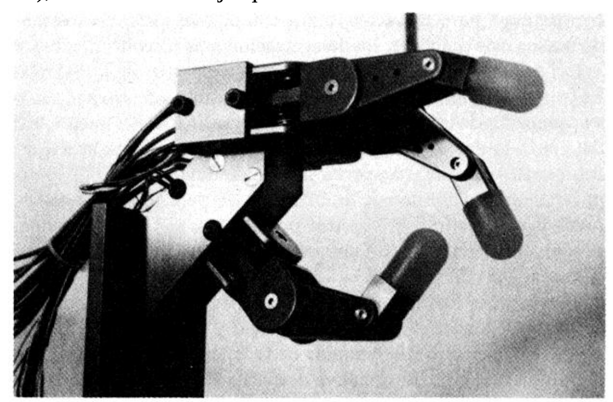
Cinco piernas
Este diseño de Hitachi consta de cinco piernas telescópicas rematadas por ruedas dirigibles. Tiene más limitaciones que el diseño con las piernas totalmente articuladas, pero se mueve mucho mejor sobre superficies planas.

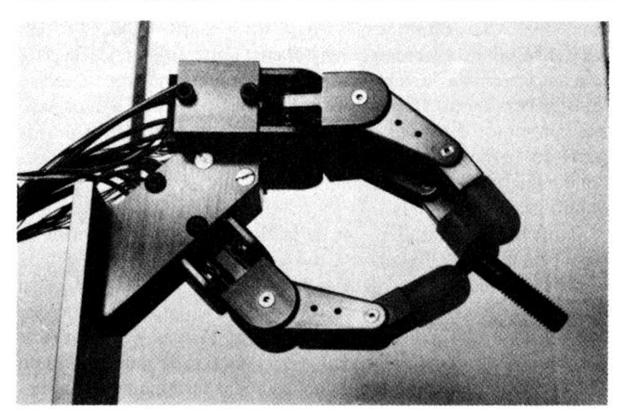
Hitachi, el gigante japonés de la electrónica, hizo experimentos a principios de los ochenta con una versión muy sencilla de la idea de los pies con ruedas. El sistema de Hitachi, para utilizar en reactores nucleares, tiene cinco «piernas» simples, verticales, como postes motorizados, que se pueden alargar y acortar. Las piernas están colocadas alrededor del robot, cada una de ellas en un vértice de un pentágono regular, y todas terminan en una rueda. Son cinco porque éste es el número mínimo de piernas que permite que el robot conserve la estabilidad cuando tiene una de ellas levantada, sin que se produzcan movimientos por el desplazamiento del peso. Las máquinas de Hitachi suben escaleras: levantan la pierna delantera hasta la altura del primer peldaño, avanzan hasta que está segura sobre él, la bajan ligeramente hasta que el contacto es firme, y repiten el proceso con la siguiente. Si las escaleras son estrechas, el robot puede tener las ruedas colocadas sobre tres peldaños consecutivos a la vez. Para salvar obstáculos, el proceso es parecido. El robot puede atravesar un suelo desigual subiendo y bajando las ruedas por la superficie de las irregularidades: una suspensión activa. La parte de arriba del robot permanece perfectamente horizontal cuando está en funcionamiento. El tiempo y las posteriores investigaciones nos dirán cuál de estas configuraciones es la más adecuada para los primeros robots «para todo».

MANIPULACIÓN EFICIENTE

Si el robot solamente se mueve de acá para allá, podrá hacer pocas tareas útiles. El trabajo productivo exige que pueda coger y transportar ingredientes, piezas, herramientas y otras cosas. Los manipuladores industriales, que son los robots más numerosos y logrados hasta la fecha, tienen brazos que llegan hasta donde sea necesario por medio de seis juntas giratorias o móviles. Si no consideramos el peso, la potencia y el control, algunos de los diseños pequeños son bastante adecuados para el robot eficiente. Como para algunas cosas es necesario poner en

contacto pares de objetos, nuestro robot tendrá, por lo menos, dos brazos. Para ciertas tareas específicas, en las que los objetos se deben poner en contacto de una manera determinada (los aficionados a la electricidad saben que una de estas tareas es soldar), sería una ventaja que tuviera un tercer brazo.





Tres dedos
La «mano de Salisbury», una solución menor para el problema de la destreza de un robot no especializado. Cada uno de los dedos va controlado por tres motores. La mano puede agarrar desde el exterior, pero tiene la misma facilidad para doblar los dedos hacia fuera para agarrar un objeto hueco.

Las manos de los robots no están tan perfeccionadas como los brazos. Los manipuladores industriales agarran y sostienen con unos accesorios fijos especiales para objetos concretos o con una especie de mano simple, de dos dedos, que recibe el nombre de *agarrador de mordaza paralela*. Son de funcionamiento sencillo, pero sólo agarran bien algunas clases de objetos ásperos. No pueden controlar ni cambiar la orientación de lo que han cogido; nuestro robot universal necesita más flexibilidad.

En algunos proyectos se han examinado los agarradores con múltiples dedos, que tienen mucha más soltura y habilidad. Uno de los mejores es el resultado de diez años de trabajo de Ken Salisbury, ahora en el MIT. Su mano de robot consta de tres dedos y puede coger y orientar huevos y tornillos y maneja cuerdas como las personas. Salisbury determinó las dimensiones de la mano y su configuración con la colaboración de un ordenador. Investigó las distintas formas de unión para encontrar el juego mínimo que permitía que las yemas de los dedos convergieran y agarraran objetos pequeños de forma arbitraria. El resultado es una mano que tiene tres dedos simétricamente colocados y que se doblan casi como los de los seres humanos. Sin embargo, como estos dedos se pueden doblar tanto hacia dentro como hacia fuera, la mano puede agarrar objetos huecos desde el interior y desde el exterior. Las fuerzas impulsoras se transmiten por medio de unos delgados cables de acero desde un conjunto de motores, que están situados a cierta distancia de la muñeca del robot.

Para que sean eficientes, las manos deben «sentir» el objeto que agarran. Salisbury está perfeccionando unas yemas de dedos hemisféricas que, por medio de indicadores de tensión cuidadosamente colocados en el interior, sientan la magnitud y la dirección de las fuerzas externas. Los programas de ordenador proyectan y ponen en práctica los movimientos de la mano y del brazo, ya que los manipuladores complejos están todavía en mantillas. Existen varios programas que proyectan movimientos del brazo complejos y libres de colisiones entre dos

puntos. Estos programas consideran un espacio (el llamado *espacio de configuración)* que describe todas las posturas posibles del manipulador. Cada articulación le añade una dimensión a este espacio y, por tanto, un manipulador tan complicado como el de Salisbury tiene un espacio de configuración complejo. Un espacio complejo significa que la búsqueda del buen camino es cara y lleva tiempo. Los tiempos de funcionamiento van desde algunos minutos hasta horas, pero se sigue investigando para conseguir mejores algoritmos, y los ordenadores son cada vez más veloces.

NAVEGACIÓN EFICIENTE

La capacidad mecánica de moverse es sólo una parte del problema de la movilidad. También hay que saber encontrar lugares específicos, volver a ellos y evitar los peligros del camino. Me he pasado la mayor parte de mi vida profesional trabajando sobre este tema, y me hace feliz poder decir que se están dando buenas soluciones. Mi tesis, que hice en Stanford en los años setenta, versaba sobre programas que tenían como objetivo que el Carro de Stanford supiera moverse por habitaciones atestadas de objetos y por espacios al aire libre. Para la primera versión de este programa (en 1976), la visión del mundo se conseguía por medio de una cámara de televisión que se había montado en el robot. El programa situaba zonas características en la imagen de televisión y las vigilaba a medida que el robot se iba moviendo. Se podía calcular su distancia y el grado de movimiento. También construía un mapa tridimensional y poco denso de los alrededores, identificaba los obstáculos y trazaba un camino hasta el punto de destino que los evitaba. El programa, entonces, movía al robot un metro aproximadamente por ese camino, miraba, organizaba, planificaba y movía de nuevo. Repetía estos cautelosos avances hasta que el Carro llegaba a su destino sano y salvo.

Por desgracia, el programa no funcionó. Un avance de cada cuatro, es decir, la parte del programa que intentaba calcular el

movimiento del robot a partir del cambio de la imagen, estaba equivocada, había identificado erróneamente las zonas que había que vigilar, informaba de un movimiento del robot que era incorrecto, y desordenaba el mapa que se iba construyendo lentamente. La probabilidad de que cruzara sin incidentes una habitación grande, un viaje de unos treinta movimientos, era tan pequeña que se podía considerar despreciable. En 1979 lo intenté otra vez con un nuevo programa y con la ayuda del nuevo *hardware*, un mecanismo que movía la cámara con toda precisión de un lado a otro del camino. Gracias a él, el programa principal podía obtener varias imágenes de la escena sin mover el robot, lo mismo que el hombre obtiene dos imágenes, una de cada ojo. Se aprovechó cuidadosamente toda la información adicional para reducir el número de errores y el programa mejoró el porcentaje de éxitos, para un solo movimiento, a casi el cien por cien. El robot solía realizar los treinta avances necesarios para cruzar una habitación y llegar a un cierto punto de destino, y mostraba un mapa correcto en una pantalla. Sin embargo, fallaba una vez de cada cuatro, bien porque al eliminar un error se había quitado del mapa un obstáculo real y el robot había chocado con él, bien porque, a pesar de todo, se habían deslizado algunos errores y le habían dado al robot una idea equivocada sobre su posición. El resultado no estaba mal para mi tesis, pero no servía para hacer que un robot ejecutara tareas complejas que le exigirían, como mínimo, que cruzara la habitación varias veces.

En 1980 me cambié a la Carnegie Mellon University para seguir con mis investigaciones en su nuevo Instituto de Robótica. Dos de los estudiantes graduados, Chuck Thorpe y Larry Matthies, examinaron y mejoraron mucho el programa antiguo e incrementaron tanto su velocidad como su exactitud. Cuando todo iba bien, daba la posición del robot (uno nuevo, que se llamaba Neptuno) con una variación de centímetros. Por desgracia, las cosas no iban siempre bien, y el porcentaje de fallos se negó tercamente a que lo modificaran. El robot cruzaba la habitación sin problemas tres veces de cada cuatro.

En 1984 nuestro grupo aceptó hacer algunas investigaciones para una empresa nueva, la Denning Mobile Robotics, Inc., de Massachusetts, que estaba creando un robot guardia de seguridad o, en otras palabras, una alarma antirrobo. En vez de una cámara, el robot llevaba una especie de cinturón, un anillo de sensores de la escala del sonar como los de las máquinas Polaroid. Ya se había comprobado anteriormente que resultaban muy útiles para detectar la presencia y la situación de obstáculos próximos, lo que permitía al robot evitarlos. Sin embargo, nuestro objetivo era mucho más ambicioso. En vez de limitarse a percibir colisiones inminentes, ¿se podría utilizar el sistema de sonar, que estaba activo continuamente, para trazar un mapa de los alrededores con el que poder dirigir la navegación punto por punto, lo mismo que lo hacían (tres veces de cada cuatro) nuestros programas guiados por la visión? Cada una de las unidades de sonar emitía un chirrido de sonido ultrasónico en un amplio cono, e informaba sobre el momento en que se escuchaba el primer eco. Este tiempo es proporcional a la distancia del primer objeto situado en el interior del cono. La distancia del objeto se puede determinar con un margen de error inferior a un centímetro, pero como el cono abarcaba un ángulo de unos 30°, la posición seguía siendo indeterminada.

Era muy distinto de las medidas tan precisas que proporcionaban las cámaras de televisión, así que los métodos que se elaboraron para el Carro no se podían usar.

Aunque una sola lectura de sonar le diga al programa muy poco sobre la posición del objeto que produjo el eco, representa un gran volumen de espacio vacío delante de ese objeto. Cuando se solapan las lecturas de los diferentes sensores, la región vacía que indica una de las lecturas puede restringir la posibilidad de localizar el objeto que ha producido el eco indicado por la otra. Si se reúnen cientos o miles de lecturas de distintas posiciones, se pueden elaborar mapas detallados que sustituirán a los sensores individuales. Como la sensibilidad de un sensor de sonar decrece suavemente desde el centro hasta los bordes de su cono, nos pareció que lo más adecuado era recurrir al cálculo de probabilidades. Alberto Elfes, otro de los

estudiantes graduados, y yo elaboramos un programa basándonos en estas ideas, y nos quedamos atónitos cuando comprobamos que conducía al robot con mucha más seguridad que el antiguo programa guiado por televisión. Bruno Serrey, otro de los estudiantes, y Larry Matthies descubrieron la manera de utilizar el cálculo de probabilidades para aplicárselo a los datos proporcionados por la televisión, y de nuevo descubrieron que funcionaba mucho mejor que el método antiguo.



Navegación autónoma
El Centinela Denning es un producto comercial que puede hacer la ronda todas las noches, durante meses y sin intervención humana, por un almacén grande o un complejo de oficinas. Va dirigido por haces emisores de luz y una imagen sonar de sus alrededores. De día, se recarga en una cabina especial.

Nuestro nuevo método representa el espacio que rodea al robot como una rejilla de celdas, cada una de las cuales contiene la probabilidad, calculada en base a todas las lecturas posibles del sensor, de que la celda correspondiente del espacio esté ocupada por materia. Una lectura puede rebajar la probabilidad de algunas celdas (por ejemplo, las que estén en el interior de un cono de sonar) y elevar las de otras (como las que están en la superficie de alcance de las lecturas del sonar). Esto nos facilitó una forma adecuada para combinar los resultados de distintos tipos de lecturas. Además, recientemente Elfes y Matthies verificaron un programa que trazaba mapas a partir de los datos proporcionados por el sonar y por la televisión. Con estos nuevos métodos, nuestros robots pueden recorrer largas distancias casi sin tropiezos. Con una nueva base matemática y con un robot nuevo y fantástico, Uranus, para seguir adelante con el trabajo, tengo una confianza casi absoluta en que la navegación se habrá dominado en el plazo de diez años, para cuando esté entre nosotros el robot universal.

RECONOCIMIENTO EFICIENTE

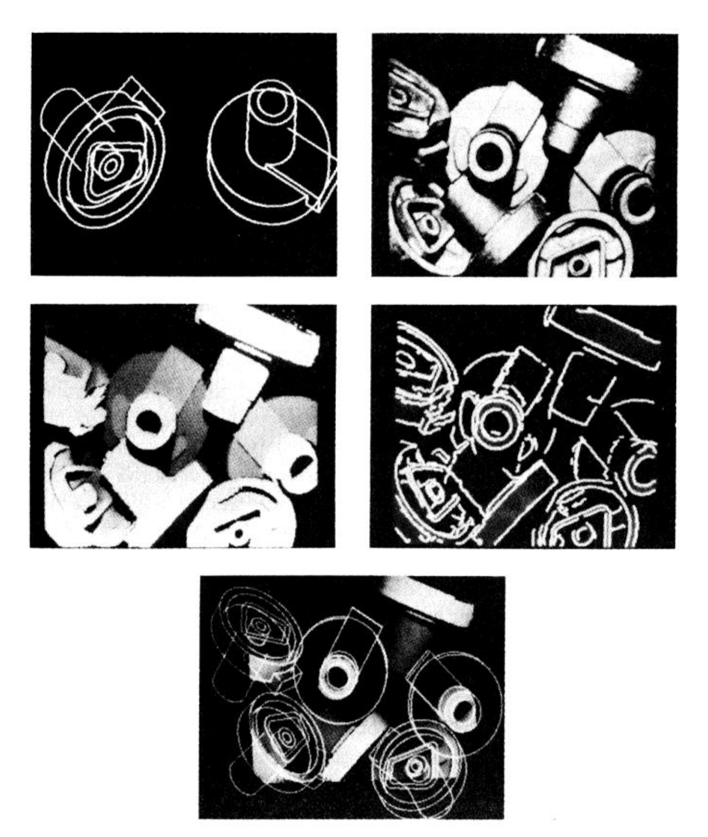
El sistema sensorial tiene otra función fundamental: el reconocimiento y la localización de objetos concretos situados en las proximidades del robot. Esos objetos que se reconocen pueden ser las pequeñas cosas que después tendrá que coger una de las manos, o los objetos grandes que sirven de señal o sitúan el punto de trabajo. Imaginemos un proceso en el que los objetos se describen por su forma y por las características de su superficie, y en el que el sistema de reconocimiento del robot busca los objetos uno por uno. Una identificación de prueba se puede confirmar si se observa la escena desde otro punto. El resultado obtenido es una descripción de la posición y de la orientación del objeto, que utilizará el programa a cargo del control de los movimientos de las manos.

La visión por ordenador es, con mucho, el vehículo más prometedor para la identificación. La operación clave consiste

en identificar un objeto concreto colocado entre otros muchos. Los estudios sobre la visión que se han realizado para robots industriales han dado con soluciones parciales al llamado problema de «escoger objetos». Estos programas le permiten a un ordenador identificar, en una imagen de televisión, objetos previamente determinados y que están entre otros muchos, incluso aunque se tapen unos a otros, de forma que un manipulador los pueda ir cogiendo uno por uno. En los años setenta, un grupo de investigación de General Motors presentó un sistema que funcionaba bien si la mayor parte de las piezas se solapaban horizontalmente. Era demasiado lento y poco fiable, así que no resultaba práctico para la producción, pero sirvió para demostrar que la idea era viable. En los últimos años, varios grupos de Estados Unidos y de Japón han elaborado programas para identificar objetos simples a partir de datos tridimensionales que se obtenían por medio de una cámara enfocada sobre un cierto lugar. Esta zona se iluminaba con ciertos dispositivos especiales que generan barras o rejillas de luz. En los ordenadores contemporáneos, estos sistemas necesitan algunos minutos para realizar identificaciones bastante poco satisfactorias. Sin embargo, es probable que en un período de diez años se consigan satisfacer los requisitos mínimos que exige nuestro robot.

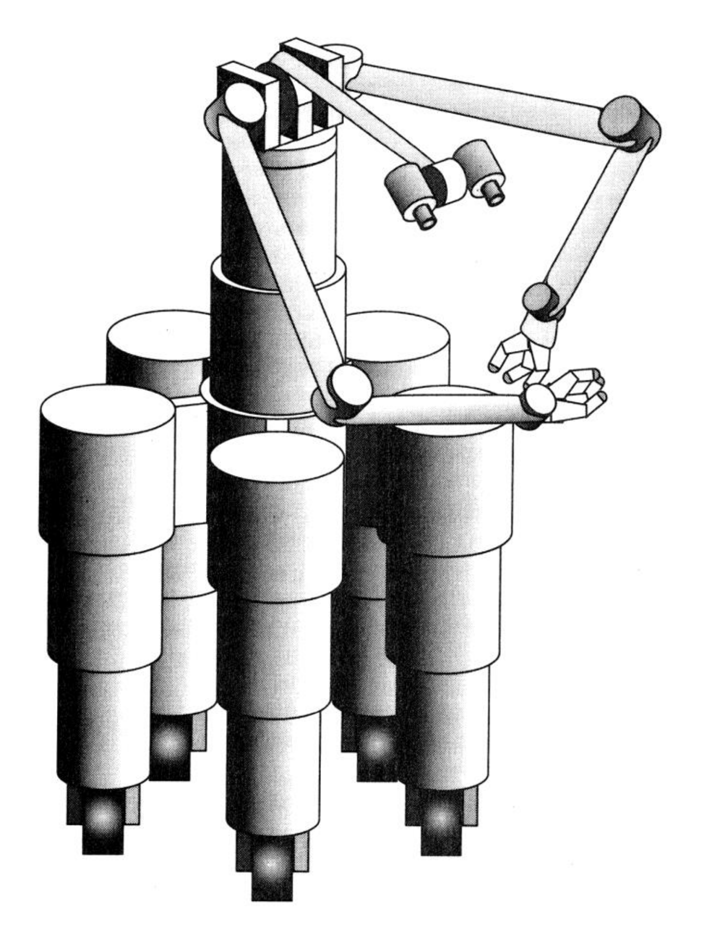
COORDINACIÓN Y PROCESAMIENTO

Los mejores prototipos para las piezas sensoriales de bajo nivel y para las que planifican los movimientos de nuestro futuro robot necesitan varios minutos de tiempo de ordenador con un microprocesador bueno. Esto, en cierto modo, sirve para medir la paciencia de los investigadores. Los procesos que están en funcionamiento durante más de una hora son, simplemente, demasiado difíciles como para poder investigarlos en profundidad, mientras que los programas más rápidos no son interesantes. Por otro lado, los tiempos de funcionamiento tienen mucho que ver con los criterios de eficiencia.



Cómo se encuentran objetos

El 3DPO (o sea, Orientación Tridimensional de las Piezas) es un programa con el cual se pueden encontrar piezas concretas que estén mezcladas desordenadamente con otras. En la secuencia de imágenes: (1) modelo tridimensional que hace el ordenador de la pieza que se desea encontrar; (2) imagen de televisión de un montón de piezas; (3) imagen del ordenador de las mismas piezas en la que el brillo indica lo cerca que está la cámara de cada porción de superficie visible; (4) el ordenador deduce los límites de superficie mayor entre el montón; y (5) el ordenador localiza la pieza modelo entre todas las demás del montón.



Un robot multiuso
Esta caricatura de un robot multiuso de la primera generación nos muestra las principales características: locomoción con capacidad limitada para subir escaleras o moverse por suelo áspero, manipulación general, visión estereoscópica, tosco sensor de 360° para moverse y evitar obstáculos. Lo que no se puede ver es ni el *software* ni el *hardware* necesarios para poner este conjunto en movimiento.

Un robot que se pase hasta una hora meditando sobre un movimiento simple es, evidentemente, inaceptable, mientras que, si el tiempo es de unos pocos segundos, es pasable. Un ordenador capaz de realizar mil millones de operaciones por segundo, con una memoria principal de mil millones de *bytes*, está bien. Ésta es, aproximadamente, la potencia de los superordenadores más grandes que se han construido hasta la fecha, y varias veces la velocidad de los mejores microordenadores. La evolución continua en el campo de los ordenadores debería hacer que este ordenador se transformara en microordenador en menos de una década. La potencia se tendrá que repartir entre algunas o muchas unidades individuales de proceso, según los avances que se realicen en las distintas líneas de evolución, y es posible que dependa de una fracción significativa del *hardware* especializado, por ejemplo, de los circuitos para los procesos de visión de bajo nivel. Pero para el tema que nos ocupa, la configuración exacta del *hardware* no es de mucha importancia.

Nuestro trabajo en la Carnegie Mellon con tareas integradas para un robot móvil sugieren que el proceso básico podría organizarse en módulos que trabajaran de forma convergente. Un programa de navegación que condujera al robot a un lugar determinado podría, por ejemplo, coexistir con otros que detectaran posibles sorpresas y peligros. Si un módulo de detección de escaleras llega a la conclusión de que el riesgo está próximo, tomaría el control del robot hasta que el peligro hubiera pasado.

UN ROBOT SENSIBLE

Vamos a ver cuál sería una de las configuraciones posibles para nuestro robot. Se mueve sobre las cinco piernas con las ruedas del diseño de Hitachi y tiene dos brazos con las manos de Salisbury. En la parte superior lleva un par de cámaras de televisión en color y una serie de sensores de sonar para percibir el mundo en las direcciones que no cubren las cámaras. También lleva un giroscopio láser barato, que le ayuda en la

navegación, y está controlado por un sistema de ordenador capaz de realizar, al menos, mil millones de operaciones por segundo. Junto con el *hardware* del ordenador, hay un sistema operativo de *software* que permite que tengan lugar simultáneamente múltiples procesos. Los programas que lleva incorporados le permiten definir los objetos identificados visualmente, solos o entre otros, y cogerlos. Al sistema de navegación se le puede pedir que almacene, elabore y compare mapas de los alrededores, que subsane los errores de los mismos y que traslade al robot a puntos específicos.

Los lectores que estén familiarizados con los ordenadores personales reconocerán la semejanza que existe con las funciones de utilidad de los sistemas operativos, especialmente con el programa de diseño gráfico *toolbox* de Apple Macintosh. El *software* es el encargado de organizar las capacidades del robot, que podrá realizar (esperamos) una cantidad asombrosa de tareas específicas. Este *software* lo proporcionarán diversos vendedores independientes. Una vez más es evidente la semejanza con los ordenadores personales. Finalmente, podemos tener la esperanza de que existan paquetes integrados de *software* que le permitan al robot pasar, de una forma rápida y automática, de una tarea a otra, con lo que este sirviente mecánico ganará en autonomía.

LA EVOLUCIÓN CONVERGENTE DE LAS EMOCIONES Y LA CONCIENCIA

Las máquinas que hemos examinado tienen un comportamiento predecible, que podemos definir como mecánico o semejante al de los insectos. ¿Seguirán siendo los robots tan predecibles a medida que se vayan haciendo más complejos, o desarrollarán algo parecido al carácter de los animales superiores y de los seres humanos?

Como hemos visto, los programas de control más avanzados de los robots móviles de la actualidad emplean datos para que los sensores mantengan las representaciones del mundo

que rodea al robot, de la posición del robot en ese mundo y de las condiciones internas del robot en diversos niveles de abstracción y de precisión. Los programas que organizan las actuaciones del robot manipulan estos modelos del mundo para crear movimientos futuros alternativos. Estos modelos del mundo se pueden almacenar y examinar posteriormente, para que sirvan de base de aprendizaje.

Un interfaz verbal que correspondiera a estos programas respondería correctamente a preguntas como «¿Dónde estás?» («Estoy en una zona de unos veinte metros cuadrados, con tres lados cerrados, y tengo ante mí tres objetos pequeños») y «¿Por qué haces eso?» («Me he dado la vuelta hacia la derecha porque creo que no quepo por la abertura de la izquierda»). En nuestro laboratorio, los programas que hemos creado suelen presentar este tipo de información del mundo del robot en forma de imágenes en la pantalla de un ordenador, una ventana directa a la mente del robot. En estos modelos internos del mundo veo el principio de la conciencia en la mente de nuestras máquinas, una conciencia que, en mi opinión, evolucionará y llegará a ser comparable a la de los seres humanos.

Los biólogos especializados en los procesos evolutivos utilizan el término *evolución convergente* para describir lo que sucede cuando especies cuya relación es muy lejana desarrollan independientemente características semejantes, posiblemente en respuesta a presiones ambientales parecidas. Los ojos son un ejemplo de la evolución convergente: han evolucionado más de cuarenta veces distintas en el reino animal. Lo que se necesitaba era la presencia de células sensibles a la luz y que las presiones debidas a la selección favorecieran a los animales capaces de ver, aunque no fuera con mucha claridad en un principio. Si una función del sistema nervioso tan compleja como la visión puede evolucionar tantas veces cuando las presiones ambientales son las adecuadas, ¿qué sucede con las emociones y la conciencia? A diferencia de la visión, estas características de la mente humana no tienen una manifestación externa indiscutible y, además, carecen de definición precisa. Toda una generación de psicólogos dedicados a estudiar el comportamiento

ha cuestionado su existencia, tanto en los animales como en los seres humanos. Sin embargo, en opinión de algunos etólogos animales, como Donald Griffin, estos conceptos son muy útiles para explicar el comportamiento animal. Si un animal actúa de la misma manera que yo cuando estoy asustado, ¿no es razonable llamar «miedo» a su estado mental? Si al enfrentarse con una situación insólita escoge entre varias alternativas complejas, igual que hago yo en las mismas circunstancias, ¿por qué no denominarlo «conciencia» en vez de decir que es un mecanismo que se llama de otra manera pero que produce el mismo efecto?

Consideremos el siguiente experimento ideal. Supongamos que deseamos fabricar un robot que haga algunas cosas corrientes como, por ejemplo, «Baja a la habitación del tercer piso, entra en ella, busca una taza y tráela aquí». Nuestra necesidad más apremiante sería encontrar el lenguaje de ordenador con el que poder definir estas tareas a la máquina móvil y un sistema de *software* y *hardware* al que incorporar ese lenguaje. Los lenguajes secuenciales de control que se utilizan en los manipuladores industriales podrían ser un buen punto de partida. Pero todos los intentos que se han hecho sobre el papel para definir las estructuras y las acciones primitivas necesarias para la movilidad revelan que la estructura de control lineal de los lenguajes más avanzados es adecuada para el brazo de un robot, pero no para una máquina que se mueve. La diferencia esencial reside en que un robot móvil, en su vagabundeo, se encuentra con acontecimientos que le «sorprenden», que no puede prever pero con los que se tiene que enfrentar. Esto implica que se activen las rutinas que controlan los sucesos posibles siguiendo un orden arbitrario y que funcionen simultáneamente, cada una de ellas con su acceso propio a los sensores, a los efectores y al estado interno de la máquina, y que exista una manera de arbitrar sus diferencias. Las prioridades de los módulos tienen que modificarse a medida que cambian las condiciones, y el control debe pasar de unos a otros.

La orden que le damos a nuestro futuro robot de que baje a la habitación del tercer piso, entre en ella, busque una taza y la

traiga, se puede llevar a cabo a través de un módulo, TRAE-TAZA, muy parecido a los de los programas escritos para los lenguajes de control de los brazos (que, además, se parecen mucho a algunos lenguajes de programación, como Algol y Basic). Existe, sin embargo, una diferencia: tiene que haber otro módulo, CUENTA- PUERTAS, que funcione al mismo tiempo que la rutina principal. El esquema para este programa podría ser parecido al siguiente:

Módulo CUENTA-PUERTAS:

Examinar el entorno del robot y comprobar si hay puertas.

Sumar uno a la variable PUERTA-NUMERO cada vez que se encuentre una puerta nueva.

Registrar la posición de la nueva puerta en la variable PUERTA-POSICIÓN.

Módulo VE-TRAE-TAZA:

Paso 1: Registrar la posición del robot en la variable POSICIÓN-INICIAL.

Paso 2: Poner la variable PUERTA-NÚMERO en cero.

Paso 3: Activar el módulo CUENTA-PUERTAS.

Paso 4: Llevar al robot paralelamente a la pared de la derecha hasta que PUERTA-NUMERO sea tres o mayor.

Paso 5: Hacer que el robot se coloque enfrente de la posición en la variable PUERTA-POSICIÓN.

Paso 6: Si el robot está ante una puerta abierta, pasar al paso 10.

Paso 7: Si el robot no está delante de una puerta, restar uno de PUERTA-NÚMERO y pasar al paso 4.

Paso 8: Si el robot está ante una puerta cerrada, intentar abrirla.

Paso 9: Si la puerta no se puede abrir, decir: «Tan, tan» y pasar al paso 6.

Paso 10: Hacer que el robot atraviese la puerta abierta.

Paso 11: Examinar los alrededores del robot y buscar tazas; si no hay ninguna, pasar al paso 15.

Paso 12: Registrar la situación de la taza más próxima en TAZA-SITUACIÓN.

Paso 13: Llevar al robot hasta que esté al alcance de TAZA-SITUACIÓN.

Paso 14: Coger la taza que está en TAZA-SITUACIÓN; si falla, pasar al paso 15.

Paso 15: Volver y colocarse enfrente de la puerta en PUERTA-SITUACIÓN.
Paso 16: Si el robot está ante una puerta cerrada, intentar abrirla.
Paso 17: Si la puerta no se puede abrir, decir: «Tan, tan» y pasar al paso 16.
Paso 18: Hacer que el robot atraviese la puerta abierta.
Paso 19: Volver a POSICIÓN-INICIAL.
Paso 20: Dejar al robot en reposo.

Hasta aquí, todo va bien. Activamos el programa y el robot, obediente, empieza a moverse por la sala contando puertas. Reconoce la primera sin error. Por desgracia, la segunda está decorada con unos carteles muy llamativos y la luz del corredor es muy mala, así que nuestro reconocedor de puertas experimental no la detecta. Sin embargo, el robot sigue avanzando paralelamente a la pared y continúa por la sala aunque en su recuento de puertas le falta una. Reconoce la puerta tres, la que se le ha ordenado que atraviese, pero cree que es la segunda y pasa de largo. Reconoce correctamente la puerta siguiente y está abierta. El programa, como cree que es la tercera, pone al robot ante ella y le hace atravesarla. Pero esta cuarta puerta da a las escaleras y el pobre robot, que no está equipado para moverse por ellas, se encuentra en peligro mortal.

Afortunadamente, en el sistema de programación que estamos usando hay otro módulo que se llama DETECTA-PRECIPICIO. Este programa funciona continuamente, examina los datos sobre el suelo que producen los procesos de visión y pide informaciones sobre el suelo de las proximidades al sonar y a los infrarrojos. Combina todos estos datos, acaso con la esperanza *a priori* de encontrar un precipicio cuando se mueve por zonas peligrosas, y proporciona un número que indica la probabilidad de que haya una bajada brusca en las proximidades. Existe otro proceso, AFRONTA-PRECIPICIO, que también funciona continuamente junto con el anterior; aunque su prioridad sea inferior, controla periódicamente este número y ajusta su prioridad basándose en él. Cuando la variable de probabilidad del precipicio (a la que podemos llamar VÉRTIGO) es bastante elevada, la prioridad de AFRONTA-PRECIPICIO

supera a la prioridad del proceso de control, VE-TRAE-TAZA en nuestro ejemplo, y AFRONTA-PRECIPICIO asume el control del robot. Un AFRONTA-PRECIPICIO bien elaborado lo que haría sería detener el movimiento del robot o frenarlo mucho, aumentar la frecuencia de las medidas del sensor sobre el precipicio, y hacer que el robot retrocediera lentamente una vez que se hubiera identificado y localizado su posición.

Pero esta secuencia de acciones tiene algo muy curioso. Si la ve una persona que no sabe nada de los mecanismos internos del robot, puede dar la siguiente explicación: «Al principio, el robot parecía decidido a atravesar la puerta, pero se dio cuenta de que había escaleras y se quedó tan nervioso y tan preocupado que se le olvidó todo lo que había hecho antes.» Como nosotros sabemos todo lo que realmente sucedió en el robot, podemos sentir la tentación de burlarnos de esta pobre persona que ha utilizado conceptos antropomórficos poco consistentes, como «decisión», «miedo», «preocupación» y «olvido», para describir las acciones de una máquina. Sí, nos puede tentar la idea de burlarnos de esa persona, pero, en mi opinión, sería un error. El robot adquirió sus manías y reaccionó con tanta honestidad como cualquier animal vivo; el comportamiento que se observó es el que corresponde a un ser que habita un mundo lleno de peligros y al que se le han proporcionado datos dudosos. Un pulpo que está buscando comida también se puede desviar de su camino ante una ligera amenaza de peligro, igual que nuestro robot. El pulpo, invertebrado, también tiene un sistema nervioso que evolucionó con independencia absoluta del de los vertebrados. Sin embargo, no tenemos ningún escrúpulo en calificar a las acciones de este animal como de «pasión», «placer», «miedo» y «dolor». Yo creo que en el comportamiento de una persona, un pulpo y un robot tenemos un ejemplo de la evolución convergente. Las necesidades de una forma de vida móvil han contribuido, en los tres casos, a crear un ser con ciertas modalidades de comportamiento para cada circunstancia, el cual modifica las mismas basándose en datos confusos, poco fiables y que pueden dar lugar a malas interpretaciones. A me-

dida que vaya aumentando la complejidad de los robots móviles, crecerá su semejanza con los animales y con los seres humanos.

Un momento, me dirá usted. Puede que haya algún parecido entre la reacción del robot y la del animal ante una situación peligrosa, pero también hay muchas diferencias. ¿No se parece más el robot a una araña asustada o incluso a una bacteria que a un ser humano que tiene miedo? ¿No reaccionaría una y otra vez exactamente de la misma manera aunque la situación no resultara peligrosa? Pues tiene razón. Creo que el sistema nervioso de la araña es una excelente analogía con los programas de robot que existen en la actualidad. Ya pasamos por la etapa de las bacterias en los años cincuenta, con las tortugas electrónicas sensibles a la luz. Sin embargo, esto no significa que haya que excluir conceptos tales como el pensamiento y la conciencia.

En su libro *Animal Thinking*, Griffin afirma que muchos comportamientos de los animales, entre ellos los de los insectos, se pueden explicar de forma sencilla en función de la conciencia: un modelo interno del ser y su entorno que, aunque sea tosco, permite considerar distintas opciones. Por ejemplo, según descubrió Otto von Frisch, las abejas pueden comunicar informaciones a otros miembros del panal sobre la orientación, la distancia y la conveniencia de un alimento por medio de la dirección, longitud y fuerza de cada movimiento de su «danza». Martin Lindauer hizo extensivas las observaciones de Frisch a los casos en los que un enjambre de una colonia con exceso de población necesita un nuevo emplazamiento. Una de las obreras del enjambre sale en busca de huecos apropiados y vuelve cuando ha encontrado y explorado uno. Entonces lleva a cabo su «danza» en la superficie del enjambre para describir la situación e idoneidad de su descubrimiento. Mientras tanto, las otras obreras «describen» otros lugares. Se envía a otras abejas a visitar y examinar los puntos más prometedores, y a la vuelta dan su «informe». Aunque a una obrera que esté «dando su informe» no le importe que otra envíe su mensaje, se la

puede «convencer» si se describe otro lugar diferente con insistencia. El debate dura varios días y se hacen repetidas visitas a un número cada vez menor de lugares posibles, hasta casi alcanzar la unanimidad. Entonces todo el enjambre se dirige al lugar elegido. Todo esto se puede explicar si postulamos que en el cerebro de cada una de las abejas existe un mapa simple en el que se describen los lugares y su idoneidad. Estos mapas se pueden modificar por la experiencia (tanto la compleja de exploración como la simple de comunicación), y se convierten en la base para la elección.

Actualmente, lo que los creadores de robots están intentando conseguir para sus robots móviles es un modelo interno del mundo que sea lo suficientemente complejo como para que pueda haber opciones en el comportamiento. A este modelo lo podemos llamar «conciencia» o no. De hecho, la investigación en el campo de la robótica es demasiado práctica como para fijarse seriamente el objetivo explícito de construir máquinas que tengan características tan nebulosas y tan controvertidas como la emoción y la conciencia. Sería suficiente con que nuestras máquinas pudieran hacer frente a todas las sorpresas, contratiempos, oportunidades, barreras y competidores que se van a encontrar en el mundo. Pero la selección natural, el mecanismo fundamental de la evolución darwiniana, es igual de utilitario con nuestros sentimientos y sentido del «yo».

En la obra *The Growth of Biological Thought*, el biólogo evolucionista Ernst Mayr afirma que tanto los sistemas vivientes como los no vivientes «casi siempre tienen la propiedad de que las características del conjunto no se pueden deducir (ni siquiera en teoría) del conocimiento completo de las partes, tanto por separado como combinadas». Con frecuencia, para explicar realidades biológicas tan complejas como la mente, la conciencia y la propia vida, se ha recurrido al proceso mediante el cual aparecen propiedades nuevas en sistemas completos. ¿Cómo pueden aparecer en nuestras máquinas algunas de las experiencias mentales más misteriosas que asociamos con los seres humanos, a la vez que perseguimos la funcionalidad? Veámoslo.

APRENDIZAJE

La babosa marina Aplysia retrae sus delicadas branquias cuando se la provoca. Si se repite la provocación, aunque sin hacerle daño, la Aplysia aprende a ignorar esa molestia y no las retrae. Si posteriormente se la somete a un estímulo más violento, como por ejemplo el contacto con un ácido fuerte, vuelve a aparecer el reflejo de retraimiento con más fuerza. En cualquier caso, es capaz de recordar durante horas el comportamiento modificado. En las últimas décadas, se ha estudiado la Aplysia con tanto detalle que se sabe perfectamente cuáles son las neuronas que intervienen en el acto reflejo. Este aprendizaje se ha identificado con los cambios químicos que tienen lugar en determinadas sinapsis de las neuronas. Si las redes de neuronas son mayores, se pueden adaptar en formas más complejas y, por ejemplo, aprender a asociar pares de estímulos específicos. Dichos mecanismos ajustan un sistema nervioso al cuerpo en que se encuentran y a su entorno. Los vertebrados deben gran parte de la flexibilidad de su comportamiento al desarrollo de esta configuración, es decir, a sistemas que se pueden activar desde distintos puntos y que fomentan o frenan las repeticiones futuras de comportamientos recientes. Aunque en los vertebrados no se conoce la arquitectura neurológica de estos sistemas, su efecto es evidente en las sensaciones subjetivas a las que denominamos placer y dolor.

En la actualidad, los sistemas de robot están configurados, en el mejor de los casos, para que sean capaces de aprender determinadas cosas de su entorno: una simple secuencia de movimientos, la localización de un componente esperado, la posición de obstáculos cercanos y, a veces, algunos parámetros para controlar un motor o para interpretar un sensor. No sirve de mucho que aprendan a coordinar sus acciones de manera complicada cuando tenemos serias dificultades en programarlos para que hagan bien una sola cosa. Sin embargo, esta etapa primitiva no durará siempre. Si comenzamos con el robot universal que describí antes, habrá que añadirle algunas funciones de aprendizaje, aunque sean de nivel muy general.

La seguridad y utilidad de un robot en el hogar aumentaría mucho si pudiera aprender a evitar determinados peligros y a aprovechar las oportunidades que se le presentan. Si una de las puertas que encuentra en su camino suele estar cerrada, merecería la pena que aprendiera a elegir caminos más largos y de menos riesgo. Una tarea resulta más efectiva si, a partir de ciertas claves, se puede aprender o incluso anticipar el punto en que se encuentra un objeto que no está siempre en el mismo lugar. Es imposible programar explícitamente al robot para todas y cada una de estas circunstancias, pero podríamos llegar mucho más lejos si contáramos con un mecanismo de acondicionamiento unificado que, por un lado, aumentara el número de decisiones que en el pasado y en ciertas condiciones habían resultado eficaces, y por otro, disminuyera las que habían tenido como consecuencia un peligro o una actividad inútil.

El *software* de acondicionamiento que tengo en mente recibiría dos clases de mensajes desde el interior del robot: unos le informarían sobre los éxitos, y otros sobre los problemas. Algunos de ellos, como por ejemplo las indicaciones de que las baterías están completamente cargadas o que cabe la posibilidad de una colisión inminente, los generaría el sistema operativo básico del robot. Otros, más específicos, destinados a que el robot realizara tareas concretas, los originarían los programas de aplicación para esas tareas. Voy a denominar «placer» a los mensajes de éxito, y «dolor» a los de fracaso. El dolor tendería a interrumpir el proceso que se estaba desarrollando, mientras que el placer aumentaría la probabilidad de que siguiera adelante.

Los mensajes, asimismo, activarían un programa que, por medio de técnicas estadísticas, «catalogaría» de manera concisa el tiempo, la posición, la actividad, el entorno y otras cuestiones que tendría que conocer el robot antes de la señal. Y también tendría que existir un mecanismo «reconocedor» que controlaría constantemente estas variables y las compararía con las entradas en el catálogo. Cada vez que se produjera un conjunto de condiciones semejantes a otras que anteriormente habían precedido al dolor (o al placer), el reconocedor enviaría

un mensaje débil de dolor (o de placer). En el caso del dolor, este mensaje de advertencia podría evitar las actividades que habían causado problemas en el pasado. Con el paso del tiempo, los mensajes de advertencia se irían acumulando en el catálogo, y el robot empezaría a evitar las circunstancias que le habían causado el problema original. Y, finalmente, la cadena de asociaciones sería tan larga que el problema se podría interceptar en una etapa muy prematura. Existen peligros, evidentemente. Si la fuerza de las advertencias secundarias no se debilita lo suficiente a medida que la cadena se va haciendo más larga, el dolor se puede convertir en una fobia y el placer en una adicción, y tanto una como la otra pueden incapacitar al robot.

Además, si se permite al robot que se adapte a su entorno de forma oportunista, el mecanismo placer-dolor se puede aprovechar mucho más directamente por medio de programas de aplicación. Supongamos que el robot tiene un dispositivo para reconocer la palabra hablada. Un módulo que, simplemente, generara una señal de placer al escuchar la palabra «bien» y una de dolor al escuchar «mal», le permitiría a un usuario modificar con toda facilidad el comportamiento del robot. Si el robot estaba dando la lata, por ejemplo, pasando la aspiradora mientras estaba ocupada una habitación, bastaría con decirle «¡mal!» unas cuantas veces para que dejara de hacerlo hasta que cambiaran las condiciones o la habitación estuviera vacía.

Un robot con *software* de acondicionamiento se podría programar para que se enseñara a sí mismo. Si una de las tareas de un programa de aplicación exigiera que un cierto tipo de envase estuviera abierto, se podría escribir una lista detallada de instrucciones en la que se especificara con claridad cómo había que coger, girar y tirar para hacerlo. Por otra parte, se puede programar un robot de una fábrica para que recoja muchos envases de los anteriores, uno después de otro y, de forma aleatoria, los empuje, gire, agite y tire de ellos hasta que se abran o se rompan. El programa de adiestramiento reconocería las dos situaciones y enviaría un mensaje de placer en un caso y una

señal de dolor en el otro, antes de que el robot pasara al envase siguiente. Gradualmente, el sistema de acondicionamiento impediría las secuencias que causaban la rotura del envase y facilitaría las otras. En el programa de aplicación final se podría insertar una versión resumida del catálogo de adiestramiento, en vez de las instrucciones explícitas, y se podría combinar con catálogos de otras secciones de la tarea desarrollada por otros robots.

Para una de estas sesiones de adiestramiento se requiere una paciencia infinita pero, aun así, puede llegar a ser exasperante. Recordarán que en el programa que describí anteriormente, aquél en el cual el robot tenía que traer una taza, si el robot se encuentra una puerta cerrada y no es capaz de abrirla, se queda allí parado y repite «tan, tan» sin parar hasta que alguien se la abre. Un robot que haya actuado así, y lo hacen muchos robots de hoy en día, no es de esperar que lo haga mucho mejor en compañía de seres humanos. Pero es interesante saber que hay muchos insectos a los que se puede enseñar a que repitan estupideces una y otra vez. Algunas avispas paralizan gusanos y los depositan en madrigueras subterráneas para proporcionar alimento a los huevos que están incubando. La avispa cava una madriguera y sella su entrada. Luego sale a cazar gusanos. Cuando regresa con la víctima, la deposita en el exterior de la madriguera, vuelve a abrir la entrada y la mete dentro. Si alguien que esté haciendo un experimento aleja un poco al gusano de la entrada de la madriguera mientras la avispa está ocupada levantando el sello, ésta se desplaza hasta donde está el gusano, lo recupera y vuelve a repetir todos los movimientos de levantar el sello, aunque la entrada ya está abierta. Si la persona en cuestión vuelve a desplazar al gusano, la avispa repite de nuevo todos sus actos. Este ciclo, aparentemente, se puede repetir infinitas veces, por lo menos hasta que la avispa o la persona a cargo del experimento queden agotadas. Se puede evitar que un robot se comporte de esta manera por medio de un módulo que genere una señal débil de dolor en cada repetición. En nuestro ejemplo, se iría evitando gradualmente que llamara a las puertas, con lo que el robot quedaría libre para

realizar otras tareas o para permanecer inactivo. El robot habría adquirido la capacidad de aburrirse.

Estos módulos, que reconocen las condiciones externas y que envían mensajes de dolor o de placer de la intensidad adecuada, le darían al robot una personalidad única. Un robot industrial peligroso y grande, dotado de un detector que enviara una señal de dolor ante la presencia de una persona, tendría miedo de la presencia humana y la probabilidad de que produjera daños sería mucho menor. Con un módulo que enviara una señal de placer al encontrar basura y otra de dolor al ver la basura de nuevo, se podría confeccionar un programa de limpieza muy creativo y agresivo en la lucha contra la suciedad.

IMAGINACIÓN

Con los robots que aprenden rápidamente se podrían utilizar programas que ofrecieran varias acciones alternativas en cada etapa de la tarea a realizar. Estas alternativas le proporcionarían al robot un amplio margen para que desarrollara su creatividad. Pero un robot con un sistema de acondicionamiento simple sería un aprendiz lento. Harían falta muchas repeticiones para que estableciera correlaciones estadísticamente significativas en el catálogo de acondicionamiento. Y algunas de las situaciones que se producen en el mundo real no permiten este método tan pausado. Un robot que tuviera que desplazarse continuamente por una vía pública y necesitara mucho tiempo para reconocer los peligros que le acecharan, lo más probable es que acabara convertido en un montón de chatarra. Un robot, o un *software,* que se adaptara con lentitud a las condiciones cambiantes o a las oportunidades que se presentan en una casa, podría perder la batalla por la supervivencia económica frente a un producto más rápido de otro fabricante.

Se puede estimular el aprendizaje si se añade otro módulo principal, un *simulador del mundo.* Incluso el sencillo robot universal que he esbozado utiliza la simulación en cierta me-

dida. Para llegar sano y salvo a su punto de destino, el programa del robot universal consulta su mapa interno de los alrededores y estudia los distintos caminos que puede seguir para descubrir el mejor. Estas consideraciones son simulaciones de las acciones hipotéticas del robot. El proceso es semejante en los casos en que el robot decide cuál es el objeto que va a elegir o cuando examina las interpretaciones posibles de lo que ve con sus cámaras. Pero cada uno de estos procedimientos es específico, imita sólo un aspecto del mundo y sólo se puede usar para una función. Supongamos que el robot está equipado con un simulador mucho más potente, que permite reproducir situaciones hipotéticas y complejas en las que participan tanto el robot como su entorno. Un programa de aplicación podría utilizar este simulador para comprobar si una acción es segura y eficaz y, en consecuencia, evitar que el robot se pusiera en peligro.

Pero las cosas se ponen más interesantes cuando los sucesos del simulador se introducen en el mecanismo de acondicionamiento. Entonces, si se produjera un desastre en el simulador (por ejemplo, una caída simulada del robot), condicionaría al robot *en la vida real* y le haría evitar el suceso simulado anterior (por ejemplo, estar en la parte superior de un tramo de escaleras simulado). De esta manera, el robot se podría preparar para enfrentarse con muchos problemas futuros si, en sus momentos libres, se dedicara a simular secuencias de acontecimientos. Estas secuencias podrían ser simples variaciones de los acontecimientos reales del día. Y el robot tendría capacidad de recordar, imaginar y soñar.

La imaginación por medio del simulador es útil solamente en el caso de que el simulador pueda hacer predicciones razonablemente exactas sobre el mundo real. Y para hacer eso es necesario tener un profundo conocimiento del mundo. Me imagino que el desarrollo competitivo de simuladores, cada vez mejores, será una parte muy importante de las investigaciones de la industria de la robótica que se realizarán en los primeros años del siglo XXI. Las empresas observarán las flaquezas de sus robots en el laboratorio y en la práctica y tratarán de adaptar

los simuladores con objeto de que puedan reproducir los aspectos del mundo más importantes para el funcionamiento del robot. Los simuladores vendrán de fábrica equipados con conocimientos genéricos, pero también tendrán que aprender las características de cada nuevo lugar. Los robots más avanzados trabajarán tanto con otros robots como con personas. Esta colaboración sería mucho más eficaz si los simuladores de estas máquinas pudieran predecir, hasta cierto punto, el comportamiento de los otros modelos. Parte de este pronóstico incluiría una aproximación al estado mental del otro modelo con el fin de anticipar sus reacciones ante distintos actos. Una vez establecido un modelo interno del estado mental de otro ser, se abrirá ante nosotros un campo nuevo y muy variado. Por ejemplo, un módulo que generara mensajes de dolor al detectar aflicción en un modelo mental del simulador, condicionaría al robot para que actuara de una manera amable. Y un robot podría descubrir que le castigaban por asignar sentimientos e intenciones «robotomórficos» erróneos a otras máquinas o incluso a seres humanos.

Es evidente que sería muy fácil programar a los robots para que cometieran delitos o realizaran tareas socialmente inaceptables, y que habría maneras legales de echarles la culpa cuando esto sucediera. Pero los robots complejos a veces se meten en problemas por sí solos. Imaginemos que un robot que lleva un simulador ha sufrido varias veces las consecuencias negativas de no haber recargado sus baterías a tiempo, y, por lo tanto, está muy «atento» al estado en que se encuentran las mismas. Supongamos que no puede entrar en casa de su dueño y que la carga de su batería empieza a disminuir. El simulador del robot repasará frenéticamente las distintas opciones en busca de una solución, un conjunto de acciones cuyo resultado sea que las baterías se recarguen.

Cuando las combinaciones de los comportamientos convencionales no le acerquen a su objetivo, el simulador pasará a contemplar posibilidades más insólitas. La casa del vecino está cerca, a lo mejor tiene la puerta abierta y con toda seguridad hay tomas de corriente en el interior. El simulador descubre

una secuencia de acciones que conduce al robot hasta los enchufes. El abandono del territorio propio, y los problemas que esto puede traer consigo, llevará asociado una señal de dolor, pero la compensan el placer y el alivio que experimenta ante la posibilidad de recargar sus baterías.

El robot realiza varias simulaciones del acto de pasar a la casa de los vecinos y, en cada una de ellas, refuerza sus condicionamientos para los pasos que tiene que dar, con lo cual se incrementa la posibilidad de que suceda. Finalmente, el condicionamiento llega a tal nivel que el robot emprende un camino que probablemente le producirá más problemas de los que le había pronosticado su imperfecto simulador. No será la primera vez que la necesidad lleve a alguien a cometer un acto desesperado.

Podríamos seguir adelante con esta especulación, y atribuir gradualmente a nuestros robots unas capacidades intelectuales semejantes a las de los seres humanos. Sin embargo, me imagino que para la época en que esto sea posible, la industria tradicional del sector de la inteligencia artificial ya venderá al por mayor robots con una capacidad intelectual excepcional. Esta industria habrá desarrollado su estrategia de arriba abajo, al mismo tiempo que se producía la evolución de los robots de abajo arriba. Es posible que este matrimonio necesite muchos años para consumarse y que surjan temas tales como de qué forma puede acceder al simulador el sistema de razonamiento para producir destellos de intuición y de qué forma puede influir este sistema de condicionamiento para poder invalidar los instintos del robot en circunstancias excepcionales. Esta unión producirá unos seres que tendrán cierto parecido con nosotros pero que, en otros aspectos, serán algo nunca visto.

II. AUMENTA LA POTENCIA

En los años setenta, cuando era un estudiante graduado, tenía la impresión de que la potencia de proceso que se dedicaba a los programas de inteligencia artificial no aumentaba con la suficiente rapidez. En 1970 trabajaba con un ordenador Digital Equipment Corporation PDP-10 de alta velocidad, que utilizaba un grupo de unas treinta personas. En 1980, mi ordenador era un DEC KL-10, cinco veces más rápido y con cinco veces más capacidad de memoria que el viejo, pero lo utilizábamos el doble de usuarios. Y, lo que es peor, tenía la impresión de que la poca potencia que sobraba la absorbían algunos de los dispositivos informáticos más caros: fantásticos lenguajes de alto nivel y de tiempo compartido, gráficos, pantallas editoras, sistemas de correo, redes informáticas y otros lujos que se habían convertido en necesidades.

Esta situación del *hardware* para ordenadores era resultado de varias causas. A consecuencia de los viajes a la Luna del Apolo y de la guerra de Vietnam, se recortó la ayuda económica para la investigación científica en las universidades, que quedaron abandonadas a su suerte con equipos antiguos. Esas mismas causas produjeron una recesión en las industrias técnicas: los ingenieros en paro abrían restaurantes baratos en vez de dedicarse a diseñar ordenadores. El impulso inicial, que tanto éxito había tenido, para resolver los problemas que se planteaban en el campo de la inteligencia artificial, todavía no se había agotado y, sin embargo, a muchas personas les parecía que las máquinas existentes ya no se podían mejorar más. Lo único que había que hacer era elaborar los programas apropiados. Pero los progresos de la investigación eran lentos, dificultosos y frustrantes, y muchos de los mejores programadores se dedicaron a la tarea, mejor remunerada, de crear herramientas de *software*, interesantes pero muy caras, cuyo éxito hizo que se crearan muchas más.

La década de los setenta fue un período de calma para el *hardware* de ordenadores, pero la de los ochenta lo compensó con creces. El salto al espacio de los rusos, en los sesenta, dio su primer impulso a la inteligencia artificial. El segundo, que tuvo lugar en la década actual, fue el salto japonés al mercado americano. El éxito industrial de los japoneses hizo que la atención del mundo entero se concentrara en la importancia que tiene la tecnología, en especial los ordenadores y la automatización, para las economías modernas. La respuesta de las industrias y del Gobierno americano fue dedicar presupuestos a la investigación. Los japoneses atizaron el fuego, bajo la influencia de un pequeño grupo de importantes investigadores, y anunciaron osadamente que se fomentarían las iniciativas para los ordenadores del futuro, el denominado proyecto Quinta Generación, que seguiría las líneas de expansión más prometedoras iniciadas por los europeos y los americanos. La respuesta americana consistió en incrementar todavía más las inversiones.

Por otro lado, con independencia del auge económico, los sistemas de circuitos integrados habían evolucionado tanto que, en los años ochenta, un ordenador completo ya cabía en un único *chip*. De pronto, los ordenadores estaban al alcance de todo el mundo, y nació una nueva generación de clientes y fabricantes de ordenadores. Por otro lado, los superordenadores, que en su momento estaban reservados a los laboratorios y agencias gubernamentales, se impusieron en cientos de industrias e institutos de investigación. Y con esta amplia gama de modelos, la industria de ordenadores llegó a ser tan lucrativa y competitiva como nunca lo había sido, y se fabricaban a ritmo frenético nuevas generaciones de máquinas más rápidas y más económicas.

¿Hasta qué punto deberá avanzar esta evolución para que nuestros ordenadores tengan la suficiente potencia como para aproximarse al intelecto humano? No es mucho lo que se sabe sobre el funcionamiento global del cerebro humano, ni tampoco sobre el funcionamiento de un ordenador inteligente, así que no podemos llegar a esta conclusión directamente. Me he

enfrentado con el problema indirectamente, y he comparado una parte del sistema nervioso que se conoce bastante bien (la retina del ojo) con los programas de visión que realizan casi la misma función. Luego efectué una extrapolación de la relación que existe entre esa comparación y el cerebro humano, a fin de obtener la cantidad de potencia de ordenador que necesitaría una máquina para imitar ese órgano. El cálculo del momento en que dispondremos de dicha máquina se realiza proyectando hacia el futuro la trayectoria de la potencia de un ordenador en relación con su coste por unidad, tal como se viene desarrollando a lo largo de este siglo.

Las cifras utilizadas para estas comparaciones provienen de mis propias investigaciones. Todo lo referente a la neurobiología ha sido extraído del libro de John Dowling, una autoridad en la materia, *The Retina,* y del ya clásico texto de Stephen Kuffler y John Nichols, *From Neuron to Brain*. Estas cifras son un tanto precarias, dado que tanto la visión del ordenador como nuestro entendimiento de la visión biológica (por no mencionar las otras funciones del cerebro) están en su infancia. Muchos de los principios básicos que rigen este complejo territorio son aún un misterio. Afortunadamente, mi comparación no requiere una precisión absoluta; un error de cien unidades arriba o abajo no implicará una diferencia cualitativa frente a las grandes escalas logarítmicas que aparecen en este capítulo. Asimismo, espero que mis errores se compensen entre sí, al menos parcialmente. En esta carretera hay muchas curvas peligrosas, así que ¡agárrense!

CIRCUITOS NEURONALES

En realidad, la retina es una extensión alargada del cerebro. Pero gracias a su situación en la parte posterior del globo ocular, a cierta distancia de la masa cerebral, es relativamente fácil estudiarla, incluso en los animales vivos. Si se la separa del cuerpo, la retina puede seguir funcionando durante horas, y todas sus entradas y salidas son accesibles. Es tan delgada y

transparente como una hoja de papel, y se puede teñir para observar determinadas neuronas por medio de microscopios ópticos y electrónicos. Por todas estas razones, la retina es la parte del sistema nervioso de los vertebrados que más se ha estudiado. Pero antes de examinarla con detalle, hablaremos de las células nerviosas.

Las neuronas, como cualquier otra célula, son unos mecanismos asombrosos. Comienzan su vida cuando se diferencian de las células de base en las primeras etapas del crecimiento del embrión, y pasan por varios ciclos de tipo ameba hasta llegar a destinos precisos dentro del cuerpo, donde se dividen y diferencian aún más. Cuando llegan a su destino final, extienden sus fibras para realizar unas conexiones específicas con otras neuronas, y estas uniones reciben el nombre de sinapsis. Las distintas subpoblaciones de neuronas se distinguen unas de otras por su tamaño, geometría y función. Algunas poseen miles de fibras minúsculas llamadas dendritas, que pueden «hospedar» a cientos de miles de sinapsis. Una de las fibras, conocida como axón, puede alcanzar una longitud de varios centímetros, un millón de veces más que el tamaño original de la célula.

Una neurona típica recibe los mensajes a través de sus dendritas y los emite a través de su axón, que puede ramificarse en la punta. Transmite estas señales a través de su membrana externa por medio de impulsos eléctricos que tienen una diferencia de potencial de algunos milivoltios. Este voltaje se mantiene constante en un ambiente húmedo favorable a la conducción eléctrica, por medio de unas «bombas» iónicas moleculares que mueven iones de potasio, sodio, calcio, cloro y otros a través de las membranas celulares. Estas «bombas» son activadas o inhibidas por unas pequeñas moléculas producidas por otras neuronas, que se transmiten por medio de distintos tipos de sinapsis. Reciben el nombre de neuronas neurotransmisoras.

Cuando una neurona recibe un impulso de un neurotransmisor, el voltaje puede aumentar o disminuir según el tipo de sinapsis o de neurona. Si hay muchas señales y, en consecuen-

cia, el voltaje disminuye por debajo de un cierto nivel, se produce una especie de cortocircuito: el voltaje sufre una caída total y repentina que se propaga a través del axón en forma de pulsación. Cuando la pulsación alcanza una sinapsis que pone en contacto una neurona con otra, desencadena una emisión del neurotransmisor desde las fosas de su membrana. Se produce una difusión a través de la sinapsis que, finalmente, hace que aumente o disminuya el voltaje de la segunda célula. Mientras tanto, las «bombas» de la primera neurona trabajan para devolver el voltaje a su nivel original, y después de unas milésimas de segundo la célula está preparada para empezar un nuevo ciclo. La intensidad del estímulo viene dada por la velocidad a la que se repiten las pulsaciones, y pueden producirse desde ninguna a varios cientos de pulsaciones por segundo. Las pulsaciones se utilizan para efectuar comunicaciones de largo alcance, pero las neuronas próximas entre sí, como las de la retina, se comunican simplemente respondiendo a los ligeros cambios de voltaje de sus compañeras. Además de las conexiones sinápticas con otras células, muchas neuronas y sinapsis cuentan con receptores para ciertos tipos de neurotransmisores «autónomos», que transportan la sangre desde otras zonas del sistema nervioso u otras partes del cuerpo, los cuales inhiben o aumentan la respuesta de la neurona.

En el núcleo de la neurona se producen procesos genéticos más lentos para fabricar los elementos neurotransmisores y transportarlos por el axón hasta las fosas de almacenamiento. La maquinaria genética de la neurona también almacena energía, crea y repara estructuras y realiza todas las demás funciones sorprendentes que son necesarias en la vida de una célula. Por suerte para los que trabajamos en la construcción de imitaciones electrónicas del sistema nervioso, dicha complejidad no participa directamente en los procesos de percepción, actuación y pensamiento. La mayor parte del mecanismo de la neurona se utiliza para el crecimiento y para crear un organismo desde dentro hacia fuera. Parece incluso como si sus operaciones para procesar la información hubieran surgido de esta necesidad evolutiva, y así nos lo demuestra.

En esta fase de la tecnología informática es más fácil mantener la maquinaria de construcción y reparación fuera que dentro. Las fábricas producen con bastante eficiencia circuitos integrados montados en el *hardware*. Esto simplifica bastante la elaboración del producto final. Además, debido a su método de funcionamiento indirecto, las neuronas son bastante lentas; parecen incapaces de generar poco más de cien señales por segundo. Hoy en día, los interruptores electrónicos, que siempre han sido muchísimo más simples que las neuronas y que ahora son más pequeños que ellas, pueden funcionar a unas velocidades que llegan a *cien mil millones* de señales por segundo. Esta gran velocidad es una de las ventajas de la electrónica, que nos permite trabajar con un número de interruptores menor que los que existen en el sistema nervioso humano. La electrónica es también excepcionalmente precisa y permite que se hagan las cosas de forma sistemática y eficaz.

Volvamos a la retina humana: ¿qué es lo que hace? Si comparamos las funciones de sus cinco tipos de células, encontraremos una respuesta un tanto improvisada pero eficaz. En su capa más externa se encuentra una red de neuronas que responden al contraste, al movimiento y a otras características propias del objeto observado. Esta red está conectada a otra capa de células capaces de detectar la luz o fotocélulas. Las fotocélulas se dividen en células cónicas, que en su conjunto pueden distinguir los colores, y células en forma de bastón, que no poseen esa capacidad.

El hecho de que la luz deba atravesar la red de neuronas para llegar hasta las fotocélulas constituye una característica propia de la retina de los vertebrados que apareció en las primeras etapas de su historia evolutiva y que permanece hasta hoy. Las retinas de los invertebrados, como el pulpo y el calamar, evolucionaron de forma independiente y los receptores de la luz se encuentran en la parte frontal de las mismas. Esta extraña posición de la red de neuronas en la retina de los vertebrados ha limitado su tamaño en gran medida, pero la fuerte presión de la selección ha realzado tanto su eficacia como su funcionamiento. Es probable que las pequeñas diferencias en

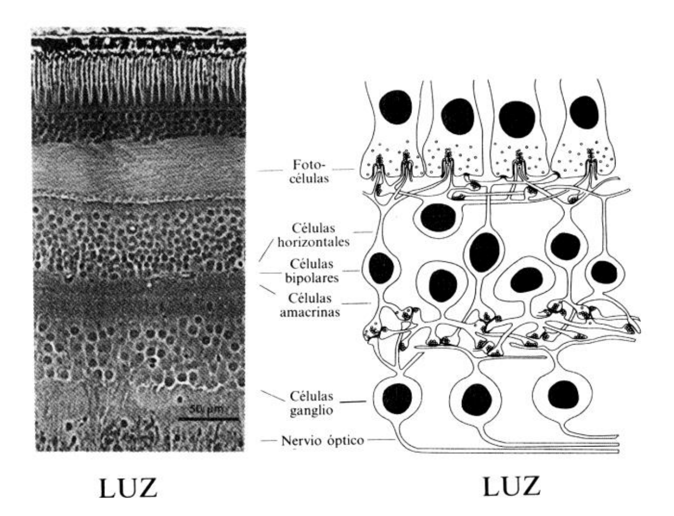
la velocidad o la agudeza visual hayan supuesto la vida o la muerte para nuestros antepasados, y las neuronas de la retina se encuentran en una posición inmejorable para abstraer de forma rápida y completa las partes esenciales de una imagen. Por lo tanto, la retina es una pieza excepcionalmente eficaz de la maquinaria neurológica de los vertebrados.

Una vez adaptados a un cierto nivel global de luz, los grupos de fotocélulas generan un voltaje que es proporcional a la cantidad de luz recibida. Hay dos tipos de células que reciben esta señal: las horizontales y las bipolares. Las células horizontales, cuyos miles de fibras recubren grandes áreas circulares de fotocélulas, producen una especie de «media» de estas áreas. Si sobre una pantalla de televisión hiciéramos un mapa de los voltajes de todas las células horizontales, veríamos una versión borrosa de la imagen de la retina. Por el contrario, las células bipolares sólo están conectadas a áreas pequeñas, y la imagen de la pantalla sería nítida. Algunas de las células bipolares pueden recibir información de las células horizontales contiguas y, en consecuencia, calcular la diferencia entre las pequeñas áreas centrales bipolares y el gran entorno horizontal. Vista en nuestra pantalla televisiva, la imagen sería más pálida que el original, con excepción de los bordes de los objetos donde se vería un halo brillante.

Los axones de las células bipolares están conectados a unas sinapsis complejas de múltiples niveles sobre las células sin axones, llamadas amacrinas. Las células de ganglio recogen información de varias de estas sinapsis de amacrinas y producen un impulso que viaja a través de su largo axón. Las células amacrinas están conectadas a células de ganglio y células bipolares, y parece que algunas conexiones envían y reciben señales.

Algunas células amacrinas realzan la respuesta «entorno central»; otras detectan cambios de luminosidad en algunas partes de la imagen. Sobre la pantalla, algunas de estas células sólo mostrarían objetos que se movieran de izquierda a derecha, mientras que otras revelarían otras direcciones de movimiento. Las células de ganglio están conectadas a varias células

bipolares y amacrinas y producen un flujo de pulsaciones cuya velocidad es proporcional a una cierta característica computada de la imagen. Algunas informan sobre el alto contraste de ciertas zonas de la imagen, otras sobre los distintos tipos de movimiento o sobre las combinaciones de contraste y movimiento.



La retina
Vista de un corte transversal, de medio milímetro de espesor, de una retina humana. El cristalino está abajo y la luz debe atravesar tres capas de circuitos que procesan la imagen para llegar a las fotocélulas, sensibles a la luz, que están en la parte de arriba. Las células horizontales calculan la intensidad media que hay en las amplias zonas ocupadas por las fotocélulas; las células bipolares restan las medias de las zonas grandes de las medias de las zonas pequeñas, obteniéndose como resultado las señales del «entorno central», que son más fuertes en los bordes del objeto. Algunas células amacrinas intensifican todavía más las señales del entorno central, mientras que otras comunican la presencia de objetos en movimiento. Las células ganglio completan los cálculos y envían los resultados, en forma de impulsos, a través de sus largos axones, a lugares situados en las profundidades del cerebro.

La televisión a la cual me refería no es totalmente imaginaria. Junto a mí, mientras escribo estas líneas, se encuentra un monitor que a menudo expone imágenes muy parecidas a las que acabo de describir. No provienen de la retina de un animal, sino del ojo de un robot. La imagen de una cámara de televisión del robot se convierte electrónicamente en una serie de números que pasan a la memoria del robot. Los programas del ordenador combinan estos números para deducir datos del entorno del robot.

Aunque se haya diseñado sin prestar gran atención a la neurobiología, muchos de los pasos de un programa son muy semejantes a las células de la retina (un caso de evolución convergente). Este paralelismo nos permite medir la potencia de cálculo del tejido neuronal.

CÉLULAS Y CICLOS

La retina humana consta de cien millones de fotocélulas, decenas de millones de células horizontales, bipolares y amacrinas, y un millón de células de ganglio, y cada una de ellas aporta al nervio óptico una fibra para transmitir señales. Todo lo anterior está encerrado en un volumen de medio milímetro de espesor por menos de un centímetro cuadrado, es decir, 1/100.000 del volumen del cerebro. Las fotocélulas interactúan con sus vecinas para aumentar su rendimiento, y parece que su elevado número es una forma de incrementar al máximo su sensibilidad: hay veces en que un simple protón puede producir una respuesta detectable. También parece que la sinapsis en las células horizontales y bipolares y en las amacrinas se realiza por medio de una computación única. Sin embargo, el límite inferior es que cada uno de los axones del millón de células de ganglio informa sobre una función específica, computada sobre una porción concreta de fotocélulas.

Para encontrar el equivalente de esta función para los ordenadores, lo primero que tenemos que hacer es equiparar el de-

talle visual del ojo humano en el equivalente de nuestro ordenador. No podemos contar simplemente las fotocélulas del ojo ya que, como funcionan en grupos, el número que obtendríamos sería excesivo. Los *tests* de agudeza visual son mejores, pero los complica el hecho de que la zona central de la retina, la *fóvea*, pequeña y densa, tenga un alto poder de resolución, diez veces superior al del resto del ojo. Aunque ocupa menos del 1 por 100 del campo visual, la fóvea utiliza la cuarta parte de los circuitos de la retina y de las fibras del nervio óptico. En condiciones visuales óptimas, esta región central tiene un poder de resolución horizontal de hasta 500 puntos distintos. Esta hazaña la puede igualar una cámara de televisión que tenga 500 elementos de imagen o *pixeles* diferentes en dirección horizontal. La resolución vertical de la fóvea es semejante, por lo que nuestra cámara necesitaría un total de 500 × 500 = 250.000 *pixeles*, que es la resolución de una imagen de buena calidad en un aparato de televisión normal.

Pero, ¿acaso nosotros no vemos mejor que la televisión convencional? No exactamente. Esta cifra de 500 × 500 corresponde solamente a nuestra fóvea, que ocupa sólo 5º de nuestro campo visual. Una pantalla de televisión estándar subtiende unos 5º cuando se la mira desde una distancia de unos diez metros. A esa distancia no son visibles ni las líneas ni otros defectos de la imagen, porque el poder de resolución de nuestro ojo no es mejor. A una distancia menor, podemos concentrar nuestra fóvea en porciones más pequeñas de la imagen y observar más detalles, lo que nos hace creer que vemos toda la pantalla así. No es cierto. Nuestros ojos se mueven de forma inconsciente y, simplemente, trasladan la zona de la fóvea de un lugar a otro de la pantalla. En algún lado, en alguna parte de nuestro cerebro que todavía permanece en el misterio, se sintetizan todas estas imágenes fragmentarias y se unen como las piezas de un rompecabezas.

De este modo, el conjunto de los circuitos de la fóvea toma una imagen de 500 × 500 y la procesa para producir 250.000 valores, de los cuales dedica parte a centrar y parte a detectar movimientos. ¿Con qué velocidad sucede esto? La respuesta

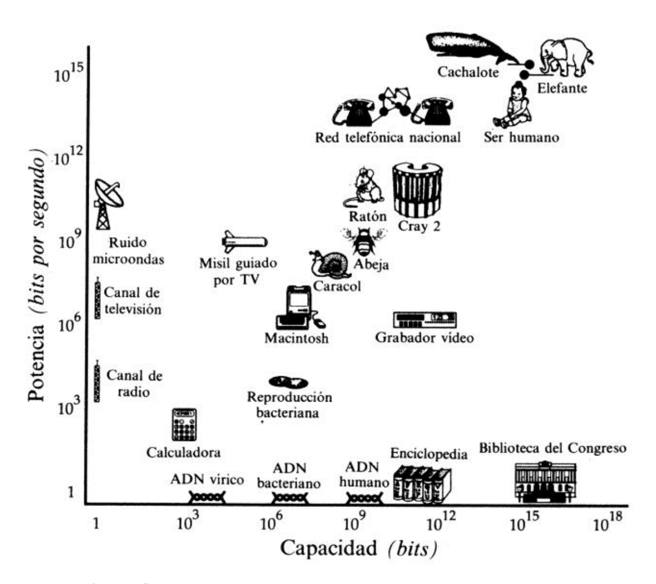
nos la dan los experimentos realizados con imágenes en movimiento. Cuando se presentan fotogramas a una velocidad menor de diez por segundo, aproximadamente, se pueden distinguir unos de otros. A velocidades superiores, se transforman poco a poco en lo que parece un movimiento lento. Aunque a velocidades superiores a diez fotogramas por segundo no se pueden distinguir unos de otros, si la luz parpadea a la misma velocidad se pueden detectar los destellos hasta que alcanzan una frecuencia de cincuenta por segundo aproximadamente. Según cabe suponer, los detectores del cambio de luminosidad se disparan en el intervalo de diez a cincuenta, pero las cadenas de neuronas más complicadas no tienen tiempo para reaccionar. Cuando la velocidad a que pasan los fotogramas es baja, los proyectores de cine evitan la mayor parte de los parpadeos por medio de un obturador que gira y que sirve para iluminar cada fotograma más de una vez. La televisión hace lo mismo y examina dos veces cada fotograma, una para las líneas pares y otra para las impares. Las zonas periféricas de la retina tienen detectores para el movimiento más rápidos que la fóvea (probablemente para observar posibles peligros que se mueven rápidamente y que vienen por los lados), y mucha gente puede detectar los parpadeos de la televisión y de las películas con el rabillo del ojo.

En nuestro laboratorio de Carnegie Mellon hemos elaborado con frecuencia programas de ordenador para realizar operaciones de centrado de imágenes que provenían de robots equipados con televisión, y hemos redactado una o dos veces programas detectores de movimientos. Para ir más rápido y realizar el trabajo con eficacia, hemos hecho un gran esfuerzo en la elaboración del programa e invertido mucho tiempo en la artillería matemática. Pero, a pesar de todo, no hemos podido llegar a la velocidad de diez fotogramas por segundo, porque nuestros ordenadores son demasiado lentos. Con un buen programa, un cálculo de centrado aplicado a cada uno de los *píxeles* de una imagen de 500 × 500 requiere unos veinticinco millones de operaciones del ordenador, lo que supone unos cien cálculos para cada valor de centrado obtenido. Con un costo

parecido, se puede elaborar un detector de movimiento. Si traducimos estos valores a términos de la retina, significa que cada célula de ganglio informa al ordenador del equivalente a cien cálculos por décima de segundo, lo que representa mil cálculos por segundo. En consecuencia, el millón de fibras del nervio óptico realiza mil millones de cálculos por segundo.

Si la actuación de la retina equivale a mil millones de cálculos de un ordenador por segundo, ¿qué diremos del cerebro en su conjunto? El cerebro tiene 1.000 veces más neuronas que la retina, pero es 100.000 veces mayor. Las neuronas de la retina, debido a las presiones del proceso de evolución, son más pequeñas y están más juntas que las otras. Si multiplicamos el equivalente de la retina en un ordenador por un coeficiente de 10.000, que sería la relación entre la complejidad cerebral y la retinal, deduciríamos precipitadamente que el trabajo total del cerebro se haría por ordenador a una velocidad de 10 billones (10^{13}) de cálculos por segundo. Esto nos daría una máquina que funcionaría a una velocidad un millón de veces superior a la de las máquinas de tamaño medio que actualmente impulsan mis robots, y sería mil veces más rápida que los mejores superordenadores de hoy.

Cifras como las anteriores están expuestas a todo tipo de críticas (ver Apéndice 1). Después de todo, la polémica surge con sólo intentar comparar ordenadores electrónicos semejantes, cuyos funcionamientos internos se comprenden y cuyas actuaciones pueden comprobarse con detalle. Por lo tanto, sería absurdo pretender que hubiera una opinión unánime cuando se trata de comparar sistemas radicalmente distintos que realizan funciones poco definidas. Sin embargo, creo que mis cifras tienen utilidad aunque sólo sean parcialmente correctas. Más adelante veremos que un error de factor mil en la relación entre las neuronas y la computación traslada la prevista llegada de máquinas totalmente inteligentes a sólo veinte años.



Potencia informática y memoria comparativas
Algunos de los organismos naturales y artificiales cuyas medidas aparecen en este capítulo. Los ordenadores que hay en los laboratorios en la actualidad son, por lo que a potencia se refiere, aproximadamente iguales al sistema nervioso de los insectos. Estas máquinas son las que han realizado todas las investigaciones en los campos de la robótica y de la inteligencia artificial. Los superordenadores más grandes de finales de la década de los ochenta equivalen al cerebro de un ratón, que pesa un gramo, pero como cuestan diez millones de dólares o más cada uno, se reservan para trabajos serios.

MEMORIA

Una vez establecido que un ordenador que realiza diez billones (10 teraflops) de operaciones por segundo es lo suficientemente potente como para albergar una mente parecida a la humana, aún hemos de decidir cuánta memoria debe tener. En 1953, el ordenador IBM 650 ejecutaba 1.000 instrucciones por segundo y poseía 1.000 «palabras» de memoria, cada una de

las cuales era capaz de almacenar un número, o una instrucción. En 1985, el Cray 2 ejecutaba un total de mil millones de instrucciones por segundo y almacenaba hasta mil millones de palabras de memoria. Esta relación, compartida por la mayoría de ordenadores, de aproximadamente una palabra de memoria por instrucción por segundo, se dio a causa de la situación del mercado y probablemente indica el tamaño necesario para almacenar problemas suficientemente grandes como para mantener ocupado un ordenador durante tiempos que van desde segundos hasta horas (a unas velocidades muy cómodas para los programadores humanos). Al existir esta relación, un ordenador «humano» necesitaría unos diez billones de palabras de memoria, o sea, unos 1015 *bits*. (Un *bit*, o dígito binario, es una pequeña unidad de información que codifica una elección entre dos posibilidades iguales. Hoy en día, las palabras del ordenador tienen entre 16 y 64 *bits*. Las máquinas más grandes suelen tener palabras más largas.)

Pero, ¿es compatible este número con lo que hoy sabemos sobre el sistema nervioso? Durante la última década, Eric Kandel, de la Universidad de Columbia, entre otros, ha estudiado los cambios celulares que tienen lugar en la Aplysia al estar condicionada por estímulos que le producen irritación. Y ha descubierto que el aprendizaje se manifiesta en forma de cambios químicos de larga duración que se producen en sinapsis individuales entre neuronas. Estos cambios afectan la fuerza de las conexiones entre las neuronas. Cada una de las sinapsis sólo podrá almacenar una de dichas fuerzas, y sólo con una precisión limitada. Si adjudicáramos a cada sinapsis un valor de diez *bits* (lo bastante como para representar un número con una precisión de tres dígitos), y si este método de almacenamiento fuera sustancialmente correcto para unos sistemas nerviosos grandes, podríamos entonces hablar de un ordenador «humano» con una memoria estándar de 1015 *bits*, capaz de contener la información codificada en las 1014 sinapsis del cerebro humano.

COMPARAR ORDENADORES

Es fácil darse cuenta de que los ordenadores son cada día más potentes, pero, ¿cuánto y a qué ritmo aumenta esa potencia? ¿Cuándo dispondremos de un ordenador de diez teraflops de un tamaño y precio adecuado para un robot autónomo? En mi primer intento de responder a esta pregunta, me pareció natural valorar los ordenadores electrónicos en cuanto a operaciones por segundo, comenzando con los primeros de los años cuarenta, para proyectar una curva hacia el futuro. Pero hubo complicaciones. Había máquinas de distintos tamaños, y sus precios oscilaban entre las decenas de dólares y las decenas de millones de dólares. Un modelo puede disponer de múltiples opciones: más memoria, procesos auxiliares, entrada y salida de datos más rápidas, etc. Las máquinas más recientes pueden ser multiprocesadoras, o sea, ordenadores múltiples que funcionan en conjunto o por separado, pero compartiendo los datos. Las máquinas poseen distintas instrucciones, y así una misma operación puede constar de diez pasos en un ordenador, y sólo de un paso en otro. Algunos ordenadores funcionan con un máximo de cinco cifras decimales, y otros pueden hacerlo con un total de veinte dígitos. Además, la literatura existente sobre las primeras máquinas electrónicas me llevó hasta sus predecesoras, los ordenadores compuestos de relés telefónicos (unos interruptores electromagnéticos que se habían perfeccionado para las centrales telefónicas). A su vez, esta investigación me llevó a unas máquinas aún más primitivas que realizaban los cálculos utilizando un motor (o incluso levas y cambios manuales). Si pudiéramos comparar estas máquinas manuales con los ordenadores automáticos, mi curva se podría remontar al siglo XIX.

Como primer paso para establecer una medida útil, decidí cancelar las diferencias de tamaño existentes entre las máquinas al dividir la potencia procesadora de cada máquina por su precio, en dólares constantes. Esto me daría una aproximación a la relación entre precio y eficacia. En cuanto a las calculadoras mecánicas, añadí al precio el valor del coste humano (o sea,

el salario del operador, valorado en 100.000 dólares de 1988), puesto que para resolver un problema una calculadora mecánica necesita de un ser humano que se dedique a dar entrada a los números y a anotar los resultados. Este planteamiento también nos permitía medir el coste aislado de un cálculo manual: un empleado sin ningún tipo de apoyo, cuyo coste efectivo en capital se sitúa en torno a los 100.000 dólares, puede realizar aproximadamente un cálculo por minuto.

El siguiente paso consistía en determinar cómo afectaban a la potencia procesadora de una máquina los factores de velocidad, tamaño de memoria, repertorio de instrucciones, etc. Esto ya era un tanto resbaladizo. A menudo, hoy en día se comparan los ordenadores en cuanto al tiempo que tardan en realizar un gran número de programas estándar. Sin embargo, no podía escoger este camino dado que la mayoría de las máquinas que pretendía incluir en mi curva ya no existían. Lo que sí sabía era cuánto tardaba la mayoría de las máquinas en sumar y multiplicar dos números, cuántas palabras de memoria podían almacenar y el tamaño de las mismas, y cuál era el tamaño aproximado del repertorio de instrucciones de cada máquina. La potencia procesadora sería el número de cálculos realizado por la máquina en un tiempo dado. Si pudiera determinar cuántos cálculos realizaba cada instrucción por término medio, para obtener la potencia total lo único que tendría que hacer era multiplicar por el número de instrucciones realizadas por unidad de tiempo. De esta manera, el problema se reducía a calcular el trabajo que hacía una sola instrucción.

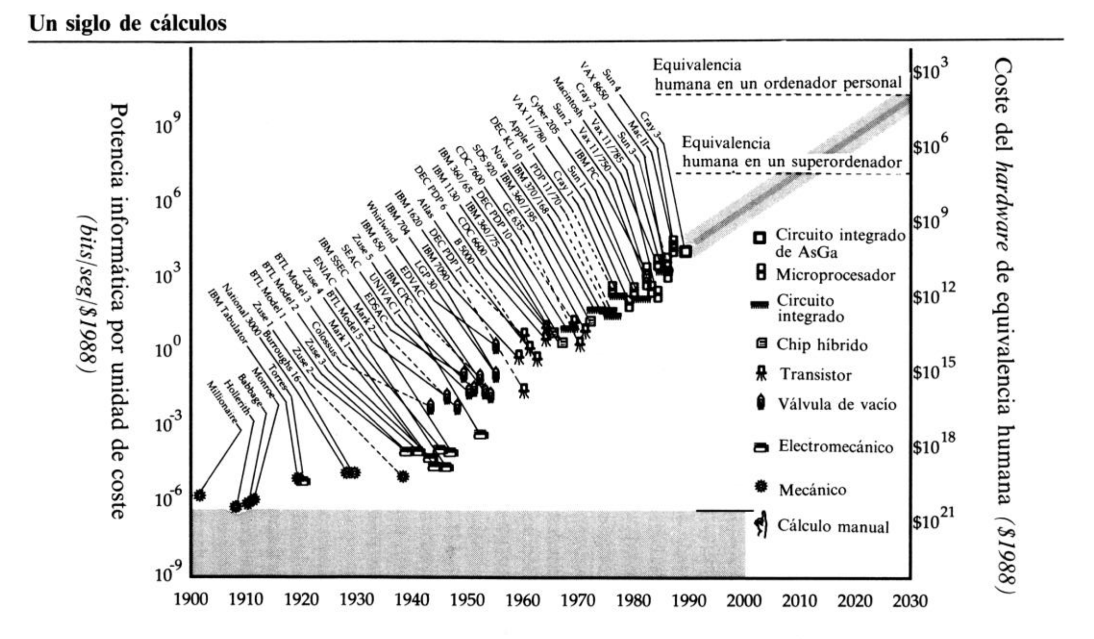
Supongamos que un cuento para niños empezara con las siguientes palabras: «Éste es mi gato. Tiene pelo. Tiene garras...» Sería muy aburrido. Imaginemos ahora otra historia que empezara así: «Éste es mi gato. Lleva sombrero. Va con una pistola...»

Mejor. La segunda historia parece más interesante e informativa, porque las dos últimas frases son menos probables: los gatos suelen tener pelo y garras, pero rara vez llevan sombrero o pistola. En 1948, Claude Shannon, del MIT, sistematizó estas observaciones por medio de un sistema matemático que recibió

el nombre de *teoría de la información*. Una de sus ideas clave es que el contenido informativo de un mensaje aumenta a medida que su probabilidad decrece. Expresado en forma matemática, es el logaritmo de la probabilidad con signo menos. Una serie de mensajes tiene un máximo contenido de información cuando su nivel de «sorpresa» es máximo.

Mi método para medir los cálculos efectivos funciona de la misma manera. Cada una de las instrucciones que ejecuta una máquina es como un mensaje. Cuanto más predecible es su secuencia de instrucciones, menos útil es el trabajo que hace la máquina. Por ejemplo, un programa según el cual un ordenador suma uno a un punto de la memoria cada millonésima de segundo no hace nada importante. Siempre se sabrá por adelantado lo que contendrá en el futuro cualquier punto de la memoria. Pero incluso los mejores programas tienen sus límites por lo que se refiere a la cantidad de «sorpresa» que pueden introducir en un cálculo en cada paso. Cada instrucción puede especificar solamente un número finito de posibles operaciones distintas y elegir entre un número finito de puntos de la memoria, cada uno de los cuales contiene un número finito de posibilidades. Estas fuentes de sorpresas se pueden combinar por medio de las fórmulas de la teoría de la información para expresar el máximo contenido de información de una única instrucción del ordenador.

En el Apéndice 2 pasaré a detallar uno de estos cálculos. Las cifras varían de una máquina a otra y de un programa a otro, pero la conclusión a la que llego es que un ordenador típico, que funcione con un programa excepcionalmente eficaz, produce aproximadamente 50 *bits* de sorpresa por operación realizada. Si el ordenador puede realizar un millón de operaciones por segundo, su potencia de cálculo es de 50 millones de *bits* por segundo aproximadamente. Utilizando estas unidades, la potencia de ordenador que precisa un robot «humano» es de unos 1014 *bits* por segundo.



Un siglo de cálculos
Coste del *hardware* de equivalencia humana *($1988)*
$10^3
$10^6
$10^9
$10^12
$10^15
$10^18
$10^21
Equivalencia humana en un ordenador personal
Equivalencia humana en un superordenador
Circuito integrado de AsGa
Microprocesador
Circuito integrado
Chip híbrido
Transistor
Válvula de vacío
Electromecánico
Mecánico
Cálculo manual
Sun 4
VAX 8650
Cray 2
Macintosh
Sun 2
Cyber 205
VAX 11/780
Apple II
DEC KL 10
Nova
SDS 920
CDC 7600
IBM 360/65
IBM 1130
DEC PDP 6
Atlas
IBM 1620
IBM 704
Whirlwind
IBM 650
Zuse 5
SEAC
IBM SSEC
ENIAC
Zuse 4
BTL Model 3
BTL Model 2
BTL Model 1
Zuse 1
National 3000
IBM Tabulator
Cray 3
Mac II
Sun 3
Vax 11/785
Vax 11/750
IBM PC
Sun 1
Cray 1
PDP 11/70
IBM 370/168
IBM 360/195
GE 635
DEC PDP 10
IBM 360/75
CDC 6600
B 5000
DEC PDP 1
IBM 7090
LGP 30
EDVAC
IBM CPC
UNIVAC I
EDSAC
Mark 2
BTL Model 5
Mark 1
Colossus
Zuse 3
Zuse 2
Burroughs 16
Torres
Monroe
Babbage
Hollerith
Millionaire
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
2020
2030
10^9
10^6
10^3
10^0
10^-3
10^-6
10^-9
Potencia informática por unidad de coste
(bits/seg/$1988)

PROYECCIONES

El gráfico anterior representa el número de *bits* por segundo de potencia que corresponde al poder adquisitivo de un dólar (de 1988) para un cierto número de máquinas calculadoras importantes, desde 1900 hasta la fecha. Aunque en los siglos XVII y XVIII se idearon y construyeron numerosas calculadoras mecánicas, solamente cuando se produjeron los avances mecánicos de la revolución industrial llegaron a ser lo suficientemente exactas y económicas como para rivalizar con los cálculos manuales. A finales del siglo xix sus ventajas habían quedado patentes, y su crecimiento fue asombroso. La escala vertical del gráfico es logarítmica: cada paso representa un aumento de diez en la relación entre rendimiento y precio. Desde 1900 se ha producido un aumento de un *billón* en la cantidad de cálculos que se pueden pagar con un dólar. Cada una de las máquinas de la figura tiene una historia fascinante, pero como este libro no es fundamentalmente histórico, me limitaré a reseñar algunos momentos culminantes. La obra *The Origins of Digital Computers: Selected Papers*, compilada por Brian Randell, contiene excelentes relatos de primera mano sobre estas primeras máquinas.

En 1834, casi con un siglo de adelanto sobre los demás, Charles Babbage, de Cambridge, Inglaterra, concibió la idea de una calculadora automática controlada por un programa. Esta «máquina analítica» era un monstruo gigantesco de vapor, con ruedas dentadas y ejes, que calculaba números decimales de hasta cincuenta dígitos. Tenía engranajes capaces de almacenar mil de estos números y una unidad de cálculo capaz de sumar dos números en menos de dos segundos y de multiplicarlos en menos de un minuto. El mecanismo de control de la máquina consistía, por una parte, en unos tambores con pernos que giraban lentamente, parecidos a los de las cajitas de música, y por otra, una sucesión de tarjetas perforadas en las que se especificaban los puntos de las memorias y las operaciones aritméticas que se podían realizar con el contenido de dichos puntos.

Teóricamente, la Máquina Analítica contaba con todos los elementos de un ordenador digital moderno. Babbage trabajó en ella los últimos treinta y siete años de su vida, pero nunca la terminó.

La gran envergadura del proyecto y el incipiente nivel tecnológico (los componentes se montaban a mano) hacían casi imposible que pudiera tener éxito. Las piezas de precisión intercambiables eran mucho más corrientes a principios del siglo XX, y en 1910 el hijo menor de Babbage presentó el funcionamiento de una parte de la unidad central de cálculo, aunque no pudo terminar la máquina. En el gráfico, la Máquina Analítica aparece como uno de los hitos de 1910, ya que es probable que se pudiera haber terminado en cualquier momento si hubiese sido necesario.

Las otras calculadoras mecánicas del gráfico funcionaban a mano y se vendían fundamentalmente a las empresas para el uso de empleados y contables, aunque algunas de ellas se utilizaron en el campo científico. Como dije antes, incluí el «precio» del calculador humano para obtener el coste de las mismas. Las primeras mejoras en velocidad y eficacia llegaron con los avances mecánicos: se perfeccionaron los engranajes y las levas de precisión fabricados en serie, los resortes y los lubricantes, y, asimismo, aumentaron tanto la experiencia en el campo del diseño como la competitividad entre los fabricantes de calculadoras. En los años veinte, las calculadoras con motores eléctricos supusieron un incremento en la velocidad y la automatización. Lo mismo sucedió cuando se incorporaron electroimanes e interruptores especiales internos por los años treinta.

En 1919, el inventor español Leonardo Torres Quevedo presentó en Madrid una calculadora electromecánica. El Aritmómetro Torres se controlaba por unas simples órdenes aritméticas introducidas por medio de un teclado de máquina de escribir, carecía de memoria y no era totalmente automático en el sentido moderno. Sin embargo, se aproximaba a ello y se podría haber adaptado para que realizara operaciones automáticas añadiéndole una unidad de grabación para la entrada de

órdenes. Con una segunda unidad para almacenar y recuperar un puñado de números habría aumentado su nivel práctico.

Por su lado, Konrad Zuse inventó en su juventud el concepto de cálculo programado, en la Alemania de 1934. En el cuarto de estar de sus padres construyó varias calculadoras electromecánicas automáticas y enormes. La tercera máquina de la serie, construida con el respaldo del Gobierno alemán, se terminó en 1941 y era ya un ordenador, controlado por una unidad de grabación con coma flotante binaria, es decir, que representaba los números en notación científica, lo que permitía cifras tanto grandes como pequeñas. Tenía una memoria de sesenta y cuatro palabras. En los años posteriores a la guerra Zuse creó una empresa que vendía modelos más perfeccionados de esta máquina.

Las máquinas de la Bell Telephone Labs (BTL) se construyeron aplicando técnicas de interruptor por relé de las centrales telefónicas. Las dos primeras fueron proyectos internos sencillos, construidos para comprobar los dispositivos de control de las armas antiaéreas. La tercera era un ordenador gigantesco, automático, multiuso, controlado por una unidad de grabación y construido con fines comerciales. Nunca tuvo éxito, y lo sobrepasaron en seguida otras máquinas electrónicas mucho más rápidas. Las colosales IBM de Harvard tenían una estructura semejante y su destino fue el mismo. Pero la era de los ordenadores multiuso de relé terminó casi antes de comenzar.

Hubo otra clase de máquinas electromecánicas con una historia más larga. De acuerdo con la Constitución de los Estados Unidos, hay que realizar un censo nacional cada diez años. Pero a medida que crece el país, cada vez se necesita más tiempo para recopilar estos datos. En 1887 se empezaron a organizar los resultados de 1880. Era evidente que si no se utilizaban técnicas más modernas, el censo de 1890 duraría hasta después de 1900. El Departamento de Censo convocó para encontrar otro sistema mejor. El ganador fue un joven ingeniero llamado Herman Hollerith, que inventó una máquina que contaba automáticamente los agujeros de unas tarjetas perforadas. En los cincuenta años siguientes, el invento de Hollerith se

convirtió en una batería de máquinas «tabuladoras» que clasificaban y ordenaban las tarjetas perforadas, las duplicaban, imprimían en y a partir de ellas, y realizaban cálculos con los datos que contenían. La empresa de Hollerith se convertiría en la futura International Business Machines (IBM), que representa en la actualidad el 70% de la industria informática.

Los ordenadores con lámparas electrónicas, en los que se aplicaban técnicas de radio y ultrarrápidas de radar, aparecieron a finales de la Segunda Guerra Mundial y fue el Gobierno el que subvencionó el proyecto. La primera máquina de este tipo que se fabricó con fines comerciales fue la UNIVAC I, y el primer comprador fue el Departamento Nacional de Censo en 1951. A finales de los años cincuenta había unos seis mil ordenadores en funcionamiento entre la industria, el Gobierno y las universidades. Sus mecanismos electrónicos se basaban en las válvulas de vacío, y se los denominó «primera generación de ordenadores».

Hacia 1960 empezó a aparecer una segunda generación de máquinas, en las que las válvulas de vacío se sustituyeron por los recién inventados transistores. Eran más pequeños, exactos y económicos, gastaban menos electricidad que los ordenadores de válvulas de vacío, y además su velocidad y su capacidad de memoria eran mucho mayores.

A finales de los años sesenta la IBM lanzó al mercado una tercera generación de máquinas que llevaban «circuitos integrados híbridos». Y se unieron docenas de diminutos transistores individuales mediante circuitos impresos sobre *chips* de cerámica del tamaño de la uña de un pulgar.

En los años siguientes, estos *chips* híbridos dejaron paso a circuitos integrados «monolíticos», en los cuales se habían grabado directamente docenas de componentes sobre *chips* de silicio de pocos milímetros cuadrados.

La tecnología de los circuitos integrados evolucionó rápidamente y, a mediados de la década de los setenta, uno de estos *chips* podía contener miles de componentes. Apareció una cuarta generación de ordenadores cuya existencia se basaba en

un puñado de estos *chips.* Pero la eclipsó rápidamente el microprocesador, un *chip* con decenas de miles de componentes que era, por sí mismo, un ordenador completo. Los avances se producían con una rapidez asombrosa y multifacética. Los ordenadores se introdujeron hasta en el sector de los electrodomésticos como, por ejemplo, los hornos de microondas, y la industria abandonó la terminología de las generaciones. El último vestigio fue el proyecto japonés de la Quinta Generación, una investigación para crear máquinas inteligentes. Escribo estas palabras con un ordenador Macintosh II, una máquina que contiene muchos *chips* con más de un millón de componentes cada uno de ellos, una máquina que no pertenece a ninguna generación.

EL EQUIVALENTE HUMANO EN CUARENTA AÑOS

Los avances del gráfico del tema **Comparar ordenadores** se han producido con sorprendente continuidad a pesar de los cambios radicales en las formas de cálculo durante este siglo. La cantidad de potencia informática que puede comprar un dólar se ha visto multiplicada por mil cada dos décadas desde el comienzo de este siglo. En ochenta años, el precio de los costes de cálculo es un billón de veces menor. Si esta velocidad continuara durante el próximo siglo, los diez teraflops que necesita un ordenador «humano» se podrían encontrar en un superordenador de diez millones de dólares antes de 2010, y en un ordenador de 1.000 dólares para el año 2030.

Pero ¿se podrá mantener esta loca carrera otros cuarenta años? ¡Con toda facilidad! La curva del gráfico no es descendente, y ya existen en nuestros laboratorios mecanismos que casi se ajustan a mis exigencias. En gran medida, la pendiente de la curva es tanto una promesa como una profecía. Los fabricantes de circuitos integrados se dieron cuenta de esta tendencia desde que, en 1963, Gordon Moore, uno de los inventores de los circuitos integrados, observó que el número de compo-

nentes de cada *chip* se duplicaba todos los años. Los fabricantes de ordenadores han hecho observaciones parecidas, y los nuevos productos de estos dos campos se diseñan sin perder de vista esta tendencia. Los fabricantes diseñan sus productos y les ponen precio para seguir la tendencia de la curva y hacer que los beneficios sean máximos. Las nuevas empresas tienen como objetivo permanecer por encima de la curva para mantener su competitividad. El propio éxito de la industria es lo que hace que ésta continúe adelante: su prosperidad cada vez mayor permite que se realicen más y mejores investigaciones, que darán lugar a posteriores avances. Y, por otro lado, los ordenadores que fabrica la industria se utilizan para diseñar los circuitos y ordenadores del futuro. Al ser cada vez mejores y más asequibles los ordenadores, también lo serán los procesos de diseño, y viceversa. La electrónica se mueve en este círculo vicioso con tanta rapidez que lo más probable es que sea la principal ocupación del ser humano para finales de siglo.

Uno de los principales factores causantes de la disminución del precio y del incremento del rendimiento es la miniaturización. Los componentes pequeños cuestan menos y funcionan con mayor rapidez. Chales Babbage lo descubrió en 1834. Escribió que la velocidad de su Máquina Analítica, que necesitaba cientos de miles de componentes mecánicos, podía aumentar proporcionalmente «si a medida que la mecánica alcanzaba estados de perfección más elevados», el tamaño de sus ruedas dentadas, que eran como la palma de la mano, se podía reducir a la escala de un reloj de pared o, mejor aún, de muñeca. Intentemos imaginar cómo sería nuestro mundo si no se hubiera descubierto la electricidad y las mejores inteligencias hubieran continuado por el camino de Babbage. Ahora existirían ordenadores mecánicos del tamaño de un escritorio y calculadoras de bolsillo con millones de ruedecitas microscópicas, que harían operaciones a miles de revoluciones por segundo.

Aunque parezca increíble, el coste por kilo de maquinaria ha permanecido constante a medida que la maquinaria se ha ido haciendo más complicada. Esto se puede aplicar tanto a los aparatos electrónicos domésticos como a los ordenadores. Las

dos categorías se han convertido en una sola en esta década. Las radios de los años treinta eran tan grandes y tan caras como los televisores de los cincuenta, los televisores en color de los setenta y los ordenadores de los ochenta. El volumen requerido para amplificar y conmutar una señal pasó de ser del tamaño de un puño en 1940 al de un dedo pulgar en 1950, al de una goma de borrar en 1960, al de un grano de sal en 1970, y al de una pequeña bacteria en 1980. En el mismo período, la velocidad básica de conmutación se multiplicó por un millón y los costos descendieron en la misma proporción. No les puedo decir con exactitud cuáles serán los avances que den lugar al coeficiente adicional de un millón que he proyectado; es imposible, por muchas razones. En el transcurso de la investigación básica se encuentran posibilidades completamente nuevas e inesperadas. Incluso entre los competidores conocidos existen muchas técnicas rivales, y se puede abandonar una línea de investigación prometedora simplemente porque surge otro enfoque que tiene una ligera ventaja. Les puedo asegurar que existen componentes experimentales en los laboratorios de hoy que son mil veces mejores que los mejores componentes comerciales, por lo menos en velocidad y tamaño. A continuación les ofrezco una pequeña lista de los que parecen más prometedores.

En los últimos años, la anchura de las conexiones de los circuitos integrados se ha reducido a menos de una micra, cifra peligrosamente próxima a la longitud de onda de la luz utilizada para imprimir el circuito. Los fabricantes han pasado de la luz visible a la luz ultravioleta de ondas cortas, pero sólo disfrutarán de un breve respiro. Los rayos X, cuyas ondas son mucho más cortas, servirían durante un período más largo, pero las fuentes convencionales de rayos X son tan débiles y difusas que precisan un tiempo de exposición largo y antieconómico. Los físicos que estudian las partículas de alta energía tienen la solución. Los electrones que se mueven a alta velocidad forman una curva en los campos magnéticos que hace que los fotones se dispersen como si se tratase del barro que sale despedido cuando gira una rueda. A este efecto se le dio el nombre

de *radiación sincrotrónica,* debido al tipo de acelerador de partículas donde se producía como interferencia. Se puede aprovechar esta energía para producir un potente haz de rayos X. Cuanto más potentes sean los imanes, más pequeño será el sincrotrón. Con unos imanes superconductores y superfríos, una máquina adecuada cabría en un camión. Sin ellos, sería del tamaño de un edificio pequeño. De cualquier manera, los sincrotrones son de gran interés y prometen reducir el tamaño de los circuitos producidos en serie del orden de las submicras. También se están utilizando los rayos de iones y electrones para «escribir» circuitos del orden de las submicras, pero los sistemas actuales sólo afectan a pequeñas regiones y de una en una, y el *chip* los debe examinar lentamente. Su propia naturaleza hace que los rayos de electrones controlados por un ordenador sean ideales para fabricar las «máscaras» que actúan como el negativo de una fotografía en cuanto a imprimir los circuitos.

Los circuitos más pequeños tienen menos inercia electrónica; cambian más rápidamente y necesitan menos voltios y menos potencia. Por el lado negativo, y a medida que disminuye el número de electrones de una señal, el circuito tendrá más tendencia a la agitación térmica. Se puede contrarrestar este efecto por medio del enfriamiento. Y, de hecho, los circuitos rápidos experimentales de los laboratorios funcionan en nitrógeno líquido superenfriado. Se está diseñando un ordenador que funcione de esta manera. El nitrógeno líquido se produce en enormes cantidades al obtener oxígeno líquido del aire, y es económico (al contrario que el helio líquido). A medida que disminuye el tamaño de los circuitos, se producen valores erráticos de los componentes como resultado del agrupamiento desigual de las principales impurezas. Por tanto, se están desarrollando nuevos métodos más precisos para implantar los nuevos circuitos. Los efectos cuánticos se hacen más pronunciados, planteándose así nuevos problemas y nuevas posibilidades. Un ejemplo de dichas posibilidades serían las superrejas, capas múltiples de regiones del espesor de un átomo con distintas densidades de silicio, hechas con haces moleculares. Éstas per-

miten afinar las características electrónicas del material, y también dan lugar a nuevos métodos de conmutación que pueden producir importantes mejoras. Aún más interesantes son los mecanismos de «punto cuántico», que aprovechan el comportamiento ondulatorio del reducido número de electrones atrapados en regiones más pequeñas que la longitud de onda del electrón.

Los primeros transistores eran de germanio; no soportaban las altas temperaturas y eran poco fiables. Al aumentar los conocimientos sobre la física de semiconductores y sobre los métodos para producir cristales de silicio, se construyeron circuitos integrados y transistores de silicio mucho más fiables y rápidos. Los nuevos materiales están alcanzando en estos momentos su «mayoría de edad». El más inmediato es el arseniuro de galio, cuya reja de cristal impide el paso de los electrones en menor medida que la del silicio, y hace que los circuitos sean diez veces más veloces. El superordenador Cray 3 utiliza unos circuitos integrados de arseniuro de galio dispuestos en un volumen de un pie cúbico con el fin de multiplicar por diez la velocidad del Cray 2. Otros compuestos, como el fósforo de indio y el carburo de silicio, están aún a la espera. El carbono puro, en forma de diamante, es una posibilidad muy buena y supondrá una mejora sobre el arseniuro de galio, como lo fue el cristal sobre el silicio. Entre sus muchas características superlativas, el diamante perfecto es el mejor conductor de calor, una propiedad muy importante en el mundo de los circuitos. La idea de un circuito tridimensional ultradenso en un diamante de alta calidad es decisiva. Hasta la fecha no se conoce ningún circuito de diamantes en funcionamiento, pero la probabilidad de que esto ocurra aumentó en 1987 después de los informes de la Unión Soviética, Japón, y con más retraso Estados Unidos, sobre la existencia de unas capas de diamantes de un espesor de hasta un milímetro producidas al calentar metano por microondas.

Es muy probable que los circuitos del futuro sean unos mecanismos cuánticos superconductores no sólo muy rápidos,

sino también muy eficaces. En los últimos veinte años, los circuitos superconductores han estado tanto de moda como pasados de moda. Han pasado por períodos de difícil aceptación debido a que el ambiente de helio líquido que requieren es muy costoso, los ciclos de enfriamiento/calentamiento son muy estresantes, y los semiconductores, que han experimentado continuas mejoras, siempre resultaban más rentables. Con el descubrimiento de los superconductores de cerámica de alta temperatura, se resolvieron muchos de estos problemas. Un transistor superconductor, presentado por los Laboratorios Bell en 1988, medía 1/20 de micra y era capaz de alternar entre encendido y apagado en un picosegundo, o sea, una billonésima parte de un segundo, partiendo de una señal de un solo electrón. Mil microprocesadores provistos de estos mecanismos cabrían en el espacio que hoy día ocupa un *chip* microprocesador, y cada uno de ellos sería mil veces más rápido. Mil procesadores, cada uno de los cuales fuera mil veces más rápido que los actuales, nos darían los diez teraflops necesarios para el equivalente humano.

Los circuitos ópticos que utilizan rayos láser y efectos ópticos no lineales para alternar la luz en vez de la electricidad, son menos conocidos. Aunque se ha podido demostrar que actúan en unos pocos picosegundos, o sea a una velocidad cien veces mayor que la de los circuitos convencionales, en la práctica sigue habiendo muchos problemas. También se han utilizado rayos láser muy finos con cristales sensibles y moléculas orgánicas para desarrollar memorias que pueden almacenar hasta un billón de *bits* por centímetro cuadrado.

En el campo de la física básica, los métodos experimentaron también avances espectaculares, que sirvieron como base a los anteriores y los precedieron. Un mecanismo no muy esperado y con pocas probabilidades es el microscopio «perforador», de precio asequible, que es capaz de ver, identificar y manipular átomos simples sobre superficies, examinándolas con una aguja afilada. La punta de la misma está sujeta por medio de tres cristales piezoeléctricos que se estiran microscópicamente bajo la influencia de pequeños voltajes. Al controlar

la corriente que de forma cuantomecánica pasa por el dispositivo, se logra un espaciado de pocos átomos de tamaño. A nivel atómico, el microscopio «perforador» supone un buen punto de partida, y sugiere grandes ideas sobre los átomos diminutos a los investigadores en los campos de los semiconductores y de la biotecnología.

Los organismos vivos son máquinas perfectas cuando se observan a nivel molecular: la información codificada en *casetes* de ARN se encargará de dirigir los mecanismos de formación de las proteínas, llamados ribosomas, que a su vez tienen la tarea de seleccionar determinados aminoácidos del entorno y conectarlos al final de las cadenas de proteínas. Las proteínas, a su vez, se desdoblan, según la secuencia de sus aminoácidos, para desarrollar muchas funciones. Algunas proteínas poseen partes móviles, parecidas a bisagras, resortes o pestillos. Otras son fundamentalmente estructurales, como los ladrillos, cuerdas o cables. Las proteínas de los tejidos musculares funcionan como émbolos de trinquete.

La industria biotecnológica actual depende de las modestas manipulaciones de la maquinaria genética natural. Sin embargo, los visionarios tienen planes mucho más elaborados: nada menos que la fusión de las técnicas biológicas, microelectrónicas y micro- mecánicas en una única e inmensamente poderosa nueva tecnología. Las técnicas de construcción por ordenador están avanzando tanto que, poco a poco, permitirán que se diseñen nuevas proteínas y se comprueben en pantallas de ordenador, de la misma manera que en la actualidad se diseñan las piezas convencionales de las máquinas. Las proteínas desarrolladas a partir de esta ingeniería, así como los mecanismos de proteínas existentes copiados de las células vivas, pueden dar lugar a unas máquinas artificiales muy pequeñas. Los primeros productos podrían ser medicinas de elaboración simple o pequeños circuitos experimentales para ordenadores. Aunque, con el tiempo, tanto las herramientas acumuladas como la experiencia misma nos permitirán construir máquinas mucho más elaboradas que, finalmente, serán tan complejas

como los pequeños brazos de robot y los, también muy pequeños, ordenadores que las controlarán. Los brazos de robot tendrían que ser de un tamaño tan reducido como para coger moléculas individuales y mantenerlas en su sitio. Y, a su vez, los robots de proteína podrían servir como herramientas, combinando átomos y moléculas de diversos tipos, para construir una segunda generación de mecanismos más pequeños, duros y resistentes. Por ejemplo, se podría disponer un cierto número de átomos de carbono formando una especie de «muro de ladrillos» para obtener un diamante perfecto con unas fibras ultrafuertes. A esta técnica se le ha dado el nombre de *nano tecnología*, dado el tamaño nanométrico de sus partes. Por el contrario, la microtecnología actual de los circuitos integrados posee estructuras de tamaño micrométrico, o sea mil veces mayores. Algunas cosas son más fáciles a escala nanométrica. Los átomos, de tamaño y forma bastante regular (aunque un tanto borrosos), se comportan de manera predecible, lo que no sucede con las piezas melladas, torcidas y cascadas de las grandes máquinas. A escala nanométrica, el mundo cuenta con una gran abundancia de componentes de absoluta precisión.

La maquinaria a escala atómica es un concepto maravilloso que nos llevaría mucho más allá de la era de los ordenadores «humanos», puesto que donde actualmente cabe un solo procesador por *chip*, habría lugar para muchos millones de procesadores. ¿Qué velocidad alcanzaría cada nanoordenador? Según la mecánica cuántica, hace falta un mínimo de energía para localizar un suceso en un tiempo dado: *energía = h/tiempo*, donde *h* es la constante básica de Planck de la mecánica cuántica. Cuanto más alta sea la velocidad, se necesitará más energía. Por encima de la frecuencia de la luz, es decir, unas 1015 transiciones por segundo, la energía llega a ser de un electronvoltio, cantidad próxima a la energía de los enlaces químicos de la materia sólida. Las velocidades mayores destruirían los interruptores. Una mil billonésima parte de un segundo, o sea un femtosegundo, es un millón de veces más rápido que el tiempo de nanosegundos existente en los componentes más rápidos de los ordenadores comerciales actuales. Por tanto, un

solo ordenador nanométrico tendría una velocidad de proceso de un billón de operaciones por segundo. Si se amontonaran millones de estos procesadores en un *chip* del tamaño de una uña, mi criterio de «equivalencia humana» sería un millón de veces más perfecto. Les parecerá suficiente, pero no puedo dejar de pensar en la, quizá posible, existencia de velocidades superiores a esta «barrera de luz».

El mundo de la física es muy turbulento y los teóricos persiguen una meta que eludió el mismo Einstein: una única teoría que englobe a todos los tipos de partículas y energías que existen en la naturaleza. En la actualidad, lo más adelantado es un modelo de supercuerdas basado en unas partículas en forma de rizos diminutos enlazadas en seis dimensiones, dos más que las cuatro del espacio-tiempo. Sus interacciones posibles predicen la existencia de una serie de partículas más pesadas que las que componen los átomos, algunas de las cuales serían estables. La materia fabricada a partir de estas partículas sería entre mil y un número astronómico de veces más densa que la materia normal. En principio, esta materia ultradensa soportaría operaciones mucho más rápidas que la frecuencia de la luz. Las ventajas de esta miniaturización no tienen por qué detenerse a escala atómica. Aunque la posibilidad de que exista una materia ultradensa estable en nuestro planeta sea una mera especulación, se conoce la existencia de inmensas cantidades de materia semejante en los enormes campos gravitatorios de las enanas blancas colapsadas y las estrellas de neutrones. Es posible que algún día nuestros descendientes exploten estos cuerpos para construir máquinas que tengan un millón de millones de millones de millones de millones (o sea 10^{30}) de veces la potencia de la mente humana.

III. SIMBIOSIS

El robot que trabaje con nosotros dentro de medio siglo tendrá algunas características muy interesantes. Su capacidad de razonamiento será sorprendentemente mejor que la de los seres humanos. Incluso los diminutos sistemas actuales son mucho mejores en algunos campos. Pero sus capacidades perceptivas y motoras probablemente se podrán equiparar con las nuestras. Y, lo que es más interesante, estas personas artificiales serán intercambiables, tanto si el otro robot pertenece a su misma generación como a otra.

Pero los robots solitarios y trabajadores, aunque sean competentes, son sólo una parte de la historia. Actualmente, y todavía durante algunas décadas, los ordenadores más eficaces funcionan como herramientas en manos del hombre. Cuando la maquinaria aumente en flexibilidad e iniciativa, la colaboración entre las máquinas y los seres humanos se podrá describir como una asociación. Con el tiempo, la relación será mucho más íntima, una simbiosis en la que ya no se apreciarán las diferencias entre el socio «natural» y el «artificial». Este camino de colaboración es muy interesante por las consecuencias que tendrá, aunque importará poco a largo plazo que los seres humanos sean o no una parte íntima del desarrollo de las inteligencias artificiales.

Empezaremos nuestra exploración del camino simbiótico con un relato de sus humildes principios, las interfaces mínimas de los primeros ordenadores.

ENSAMBLADORES Y PROGRAMAS ALMACENADOS

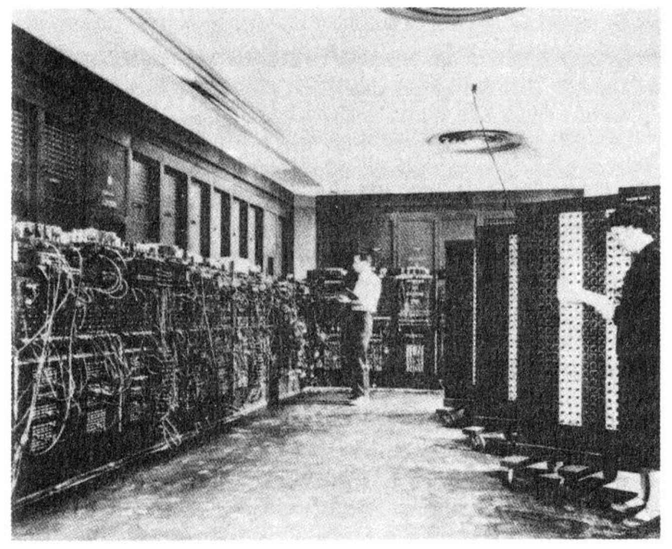
El primer ordenador digital electrónico multiuso, el ENIAC, construido en Filadelfia en 1946, fue diseñado para el control por cable. Un programa típico suponía que había que

conectar a mano miles de cables desde un punto a otro de enormes paneles de programación. Era aburridísimo «escribir» un programa de estas características. Perfeccionarlo y eliminar sus defectos era terrorífico. Antes de la famosa propuesta de John von Neumann, solamente se escribieron unos pocos. El ENIAC tenía tres grandes bloques de interruptores de disco, llamados *tableros de función,* que servían para almacenar los resultados de las operaciones matemáticas previamente realizadas y que se necesitarían durante los cálculos. Podían hacer raíces cuadradas, logaritmos y otras funciones más complicadas. Von Neumann pensó que estos interruptores se podían usar de una manera diferente, es decir, para almacenar secuencias de instrucciones, codificadas en forma de números, que dirigirían el funcionamiento de la máquina. El programa conectado permanentemente se podría elaborar de tal manera que pudiera leer las instrucciones de los tableros de función, una tras otra, y hacer lo que indicaban los números. A partir de entonces sería muy fácil programar la máquina para que realizara nuevas tareas, ya que bastaría con marcar las órdenes en los tableros de función.

Este nuevo estilo de programación era mucho más sencillo e ingenioso que el original. El programa se componía de pulcras columnas de números, ya no era como un nido de ratas lleno de cables. El sistema numérico de codificación para las operaciones de la maquina pasó a llamarse *lenguaje de máquina.* Significó un pequeño paso para los usuarios de los ordenadores, y un salto gigantesco para la organización informática.

Todos los ordenadores digitales, después del ENIAC, incorporaron una versión ampliada de esta idea del *programa almacenado.* Y no es sólo que hubiera programas representados por secuencias de números, sino que los propios números se guardaban en la misma memoria que se utilizaba para realizar los cálculos y se podían cargar a gran velocidad desde ciertos dispositivos de entrada como, por ejemplo, lectores de cintas de papel perforadas. La unidad de memoria le permitía al orde-

nador modificar su propio programa cuando estaba funcionando. Esta técnica tan misteriosa se utilizó mucho al principio, pero ahora es menos corriente. El ENIAC tenía su sede en una habitación repleta de válvulas de vacío, en la que almacenaba menos de cincuenta números y podía realizar unos mil cálculos por segundo.



ENIAC
El nido de ratas lleno de alambres que aparece a la izquierda son los tableros de programación originales de la máquina. Los tres bloques de interruptores de la derecha tenían, en un principio, la finalidad de contener las tablas de funciones matemáticas, pero pronto pasaron a ser una manera más adecuada de representar los programas.

Los sucesores del ENIAC podían almacenar programas enteros en su memoria de trabajo, porque se habían inventado nuevos métodos para almacenar, de forma más compacta y económica, las pequeñas marcas electrónicas que formaban la memoria. Algunos de ellos tenían un dispositivo parecido a un tubo de televisión que podía guardar miles de esos *bits* en forma de diminutas zonas de carga eléctrica situadas sobre su

pantalla de cristal. Las barría continuamente un haz de electrones, que podía tanto percibir como alterar el contenido de cada una de ellas. Otros tenían miles de *bits* codificados en forma de corriente de recirculación de pulsos acústicos que descendían por una larga columna de mercurio al final de la cual había un sensor electrónico que los percibía, amplificaba y volvía a inyectar en la parte superior de la columna. Otro método consistía en grabar magnéticamente los *bits* en la superficie de un tambor o un disco que giraba rápidamente. Los discos magnéticos evolucionaron y se convirtieron en los dispositivos de almacenamiento que utilizamos en la actualidad, pero eran demasiado lentos y no sobrevivieron tanto tiempo como la memoria interna de funcionamiento de los ordenadores.

El método que tuvo más éxito fue aquél en el que se empleaban unas rosquillas diminutas de material magnético fabricado especialmente que se ensartaban en las intersecciones de una red de finos alambres. Cada uno de estos *núcleos magnéticos* podía almacenar un *bit* codificado como una magnetización en el sentido de las agujas del reloj o en el contrario. La dirección de magnetización de un núcleo magnético en particular se podía modificar enviando pequeñas corrientes a través de uno de los alambres, vertical y horizontal, que lo atravesaban. Su contenido inicial se podía determinar con la ayuda de otro alambre que atravesaba en zigzag todos los núcleos. Siempre que la magnetización de uno de estos núcleos cambiaba de dirección —del sentido de las agujas del reloj al contrario o viceversa— aparecía en el alambre una pequeña pulsación. Las memorias de los núcleos magnéticos almacenaban miles y millones de números y, durante más de veinte años, fueron la forma más importante de almacenar el trabajo de los ordenadores, hasta que, a mediados de la década de los setenta, las superaron los circuitos de almacenamiento, que eran muy parecidos a los del ENIAC pero hechos de transistores en lugar de válvulas y se alineaban por miles en diminutos *chips* de silicio.

La programación en el lenguaje de máquina representó un gran ahorro de trabajo. Pero la memoria, velocidad y disponibilidad de los ordenadores, que crecían a un ritmo muy alto, empezaron a plantear a los usuarios unos problemas tan amplios y tan complejos que hasta el lenguaje de máquina se volvió insoportablemente aburrido. Un programa de lenguaje de máquina está formado por una serie de instrucciones codificadas en forma de números. El *código de funcionamiento,* que son algunos de los dígitos de cada número, especifica qué es lo que tiene que hacer el ordenador, es decir, si debe sumar o conseguir la instrucción siguiente de otro punto de la memoria. Los dígitos restantes indican la dirección de uno o más puntos de la memoria desde los que se puede, en nuestro ejemplo, encontrar los números a sumar, o pasar a la siguiente instrucción. El proceso de expresión de problemas formulados inicialmente en forma algebraica —como $x2 + y$— en códigos numéricos que pudiera utilizar el ordenador, fue lento y laborioso y era muy fácil cometer errores. Y, lo que es peor, el hecho de incluir en un programa algunas instrucciones adicionales podía suponer que hubiera que cambiar de lugar una gran parte del resto del programa y un cierto número de zonas de almacenamiento de la memoria, lo cual, a su vez, exigiría que se alteraran las direcciones de muchas, acaso miles, de instrucciones. Con que se cometiera el más mínimo error, el programa no funcionaría. A los programadores que se especializaron en esta tarea penosa y exacta, jalonada por ráfagas de invenciones artísticas, a veces se les trataba con el respeto reservado a los maestros de ajedrez. Pero no hubo manera de que fueran los propios ordenadores los que se especializaran en labor tan monótona.

A mediados de la década de los cincuenta, los programadores creaban larguísimos programas cuya función era traducir las órdenes simbólicas (como ADD X) al lenguaje de máquina, asignarles automáticamente las variables y las instrucciones y seguir su pista para saber dónde localizarlas. Los puristas del lenguaje de máquina se quejaban de que los *ensambladores,* como se denominó a los programas de este tipo, reducían el

control que uno debe ejercer sobre el funcionamiento del ordenador y que desperdiciaban mucho tiempo de ordenador haciendo la traducción. Pero con todo, los programas simbólicos representaron tal adelanto que el lenguaje escrito de máquina pasó a ser un arte en vías de extinción. El camino que unía a los seres humanos con los ordenadores se había vuelto a ensanchar, y era mucho más favorable para el tráfico cotidiano. Asimismo, facilitaba el transporte de maquinaria más pesada.

COMPILADORES Y SISTEMAS OPERATIVOS

Los ensambladores supusieron una gran ayuda para los programadores, pero seguían siendo tremendamente aburridos para los usuarios esporádicos que tenían que resolver problemas concretos. Estos consumidores exigían, y se les concedió, *lenguajes de alto nivel* en los que la notación matemática compacta, como la que usaban habitualmente en su trabajo, sustituyera a las largas secuencias del ensamblador. Por ejemplo, en un lenguaje de alto nivel como el FORTRAN, que fue uno de los primeros y sigue utilizándose, A × X + B sustituiría a la siguiente secuencia de ensamblador:

```
LOAD    A
MULT    X
ADD     B
```

Otros programas, a los que se denominó *compiladores*, auténticas demostraciones de genialidad cuando se elaboraron por primera vez, servían para expresar las líneas del lenguaje de alto nivel en las largas secuencias del lenguaje de máquina. El usuario de un lenguaje de alto nivel tenía escaso conocimiento, o ninguno, del programa final. Los compiladores consumían más tiempo de ordenador —tan valioso— que los ensambladores, y como carecían de la inteligencia y de la intuición de los programadores humanos, produjeron un lenguaje de máquina más largo y más lento que el que se producía

cuando el mismo problema se expresaba en lenguaje ensamblador. Tenía muchos inconvenientes y, hasta hace poco, muchas aplicaciones importantes exigían lenguaje ensamblador.

Pero los lenguajes de alto nivel tenían muchas ventajas para el usuario medio. Gracias a ellos, los programas eran más fáciles de escribir y se podía detectar una gran parte de los errores durante la traducción, en vez de producirse fallos misteriosos en pleno funcionamiento. Como los lenguajes de alto nivel tenían una forma muy semejante a la de la notación matemática convencional, resultaban más fáciles de aprender a los que no eran especialistas en ordenadores. Una de sus características fundamentales era la *transportabilidad*. A diferencia de los programas en lenguaje de máquina o en ensamblador, los programas de alto nivel no reflejaban el funcionamiento detallado de ninguna máquina en particular, por lo que se podían utilizar en ordenadores totalmente diferentes. Por ejemplo, existen compiladores FORTRAN para prácticamente todos los ordenadores, y algunos programas en FORTRAN han presenciado toda la historia de la informática, pasando de una generación de ordenadores a la siguiente.

En los primeros ordenadores digitales, las tareas de creación de un programa, control de su evolución y posterior limpieza se hacían manualmente. El programador podía contemplar el funcionamiento de grupos de luces que indicaban el estado interno de la máquina, interrumpirlo, examinarlo y modificarlo en cualquier momento o seguirlo paso a paso. Esto era bastante cómodo, pero muy caro, ya que el tiempo de las máquinas se calculaba que costaba cientos de dólares por hora. Para reducir el tiempo perdido, a finales de los años cincuenta las empresas de ordenadores empezaron a lanzar programas denominados *monitores, supervisores o sistemas operativos*, que servían para controlar el flujo de los sucesivos programas y leer las tarjetas perforadas o las unidades de cinta magnética, por medio de sus máquinas. Los primeros sistemas operativos seguían un método bastante absurdo. Los programas funcionaban de uno en uno; si aparecía cualquier problema, se detenía auto-

máticamente, se imprimía el contenido de la memoria y se ponía en marcha el siguiente programa. Y se le entregaba al programador el *depósito del núcleo* (como se llamaba a dicho contenido de la memoria una vez impreso) para que encontrara el problema. Muchos de los programas estaban en el ordenador solamente durante algunos segundos y los resultados que producían le ocupaban al programador hora o días, hasta que lo dejaba listo para que se volviera a probar.

Desde mediados de los años cincuenta hasta finales de los sesenta, esta forma de utilización *en serie* de los ordenadores era la regla, y los usuarios de toda una generación de los mismos, en especial los de los equipos de IBM, no conocían otra. Sin embargo, alguno de los antiguos cronometradores recordó la época en que se utilizaban las manos. Y no sólo era más sencillo encontrar un error en la programación si se tenía al ordenador como aliado, sino que también se podían escribir programas que establecían un diálogo con el usuario. Los programas interactivos hicieron posible que el usuario y el ordenador fueran como socios, y era frecuente que el usuario aportara su intuición y su buen juicio y el ordenador su memoria y su prodigiosa capacidad de cálculo. El problema consistía en establecer el nivel adecuado de servicio para que el ordenador no desperdiciara tiempo esperando a que el ser humano hiciera el siguiente movimiento.

Algunos grupos de académicos empezaron a trabajar para encontrar alguna solución, una modalidad pavorosamente compleja de sistema operativo que funcionara con varios programas a la vez, puestos en marcha desde terminales interactivas independientes. Cada uno de los programas de los usuarios funcionaría durante una fracción de segundo, luego el sistema operativo de *tiempo compartido* le pasaría el control al siguiente, y así sucesivamente hasta llegar de nuevo al primero. Un ser humano que utilizara estos programas activos sería incapaz de percibir las interrupciones y tendría la sensación de que estaba usando el ordenador él solo, aunque fuera un poco lento. Si un programa se quedaba temporalmente inactivo, en

espera de la respuesta del usuario, el sistema operativo lo ignoraría y se desperdiciaría muy poco tiempo.

Los sistemas de tiempo compartido produjeron una gran controversia en la comunidad informática, lo mismo que los ensambladores y los lenguajes de alto nivel en su momento. Como tenían que mantener en funcionamiento varios programas a la vez y decidir qué era lo que tenían que hacer varias veces por segundo, necesitaban una fracción de los recursos del ordenador mayor que la de los sistemas operativos. Los de tiempo compartido, sin embargo, tenían otras ventajas. Como había muchos usuarios activos a la vez compartiendo los recursos del sistema informático —memoria, discos, mecanismos de grabación, impresoras, pantallas, etc.— se les podía sacar mucho más rendimiento que con los sistemas de programa único. Y, lo que es más importante, los programadores podían controlar el funcionamiento de sus creaciones y hacer una parada cuando, como suele suceder con los programas nuevos, algo iba mal. Sin embargo, las mejoras más importantes, por lo que se refiere a la eficacia, fueron para los clientes, no para la maquinaria. En vez de tener que esperar durante horas para comprobar el funcionamiento de un programa, como en el caso de los sistemas en serie, los usuarios de los sistemas de tiempo compartido podían comprobar la evolución de un programa, detenerlo, modificarlo y volver a intentarlo de nuevo, todo ello en pocos minutos. Esta velocidad hizo posible que se desarrollara un estilo de programación altamente experimental, parecido al de Pavlov, que se caracterizaba por los ciclos rápidos de «castigo» y «recompensa». Había nacido una generación de hábiles expertos en ordenadores.

Estos expertos, algunos de los cuales se pasaban la mayor parte de su tiempo ante los terminales de un ordenador, empezaron a desarrollar y mejorar rápidamente las capacidades básicas de su medio favorito. Las aplicaciones completamente nuevas de los ordenadores se convirtieron en algo corriente. Hacia la mitad de la década de los sesenta, los usuarios se podían comunicar entre ellos, directamente de terminal a terminal o por correo electrónico que se podía leer y contestar cuando

se disponía de tiempo libre. Podían entretenerse con textos interactivos o con videojuegos, hacerse sofisticados trucos, compartir opiniones y programas y, en general, experimentaban una sensación de comunidad a través de las máquinas inconscientes de sus jefes. La comunidad se enriqueció cuando se crearon ciertos instrumentos de uso público, tales como archivos de ordenadores, tablones de anuncios de la comunidad, citas ingeniosas, consejos técnicos, escritos originales, programas de diversión y sistemas interactivos de información y recuperación para ayudar a clasificar todas estas posibilidades. Algunos de los sistemas más sofisticados ofrecían también dibujos y sonidos que se podían hacer con la máquina.

El uso de los ordenadores, variado e intenso, que hacían los expertos, exigía que se crearan maneras eficaces de descubrir, poner en marcha y detener muchos programas diferentes, de examinar, leer y modificar los archivos de la información, de trabajar a la vez que otros usuarios y de exigirle al ordenador que hiciera muchas cosas automáticamente. Y los lenguajes de órdenes, con los que se comunicaban los usuarios en sus terminales con los sistemas operativos de tiempo compartido, evolucionaron hasta conseguirlo. Estos sistemas, diseñados por y para expertos, con una inmensa cantidad de derivaciones que iban surgiendo sobre la marcha, y controlados por medio de poderosos conjuros que solían ser incoherentes y difíciles de recordar, exasperaban a los usuarios con menos experiencia. Para el experto, la relación con el ordenador era instintiva, rápida y muy fuerte. Nada le estaba vedado al que podía crear los conjuros apropiados en cada caso. Pero para los que no sufrían la monomanía necesaria para soportar la rápida y amorfa evolución del sistema, esta relación era algo opaco, no servía de gran ayuda y era fácil cometer errores. Los expertos encontraron la horma de su zapato cuando se conectaron sistemas muy distintos por medio de redes de ordenadores. Hasta un experto es novicio cuando desconoce los arcanos.

Uno de los grandes logros de la era de los expertos fue el sistema de tiempo compartido Unix. Lo crearon dos jóvenes a principios de la década de los setenta en los Laboratorios Bell,

y otros lo ampliaron en la Universidad de California, en Berkeley. En la década de los ochenta pasó a ser de uso corriente en los grandes ordenadores, y en la actualidad el Unix se utiliza en los ordenadores personales.

MENÚS E ICONOS

De vez en cuando, los fabricantes de ordenadores intentaban incorporar a sus sistemas, para sus clientes más serios, algunas de las innovaciones de los expertos. Uno de los objetivos más importantes fue hacer que el lenguaje que invocaba las distintas funciones del sistema operativo se pareciera lo más posible al inglés, con la esperanza de que facilitara la tarea de los usuarios que conocieran bien esa lengua. Esto ya lo había observado la comunidad científica cuando se introdujeron los lenguajes de alto nivel en los que se empleaba la notación matemática más o menos estándar. Los que estaban en contra aducían que, a diferencia de las matemáticas, el inglés es un idioma pobre, que no permite descripciones detalladas. Las comunidades en las que la precisión es esencial generan sus propias convenciones especializadas, y el ejemplo más claro es la notación matemática.

En este caso, los que se oponían tenían razón. A finales de los años sesenta y principios de los setenta se estudiaron las interfaces experimentales elaboradas en algo parecido al inglés (con largos acrónimos en los que se solían incluir las palabras *Plain* o *Simple* y *English),* pero no tuvieron mucho éxito. No era (ni es) posible expresar por medio de un programa la comprensión global que comunica un lenguaje ni el sentido común y el conocimiento del mundo que en él subyacen. Los sistemas mismos fallaron al no comprender (o malinterpretar) muchas de las frases que se les presentaban, de forma que utilizarlos se convertía con frecuencia en un juego de adivinanzas. Hacía falta mucha habilidad para expresar las peticiones en forma de frases que pudieran interpretar bien sus dispositivos para ana-

lizar el lenguaje, muy complejos pero incompletos. En comparación era muy fácil aprender un código de mandatos simple, coherente y especializado.

Las necesidades de los usuarios esporádicos de ordenadores (y de los especialistas también) estaban mejor cubiertas por la invención de muchos dispositivos simples: el *menú de múltiple elección*. En los sistemas de menú, las opciones principales se presentan en una lista y el usuario elige una de ellas. Esta elección puede llevar a un segundo menú en el que se esbozan características secundarias, y así sucesivamente hasta que quede convenientemente detallada la operación que se desea realizar. Los sistemas estándar de menú también tienen sus inconvenientes. Hace falta tiempo para que el ordenador imprima y el usuario lea los menús largos, y las posibilidades, a veces, no están expresadas de forma compatible. Por lo tanto, los sistemas de menú son más lentos que los lenguajes especializados si los emplean usuarios experimentados. Asimismo, es difícil expresar ideas realmente complejas en un juego de veinte preguntas. Los sistemas híbridos que, de vez en cuando, proporcionan respuestas en forma de ensayo, a veces resuelven este problema bastante bien.

El problema de la velocidad en los menús impresos fue fácil de solucionar, aunque un poco caro: exigía simplemente que se instalasen terminales más rápidas. El nuevo problema del usuario consistía ahora en digerir rápidamente largas listas de opciones que le resultaban desconocidas, y la solución no era tan sencilla. Los expertos que trabajaban en el lujosísimo Centro de Investigación de Xerox en Palo Alto (PARC) encontraron y desarrollaron una solución excelente. En los primeros años de la década de los setenta, este grupo organizó unos costosos centros de trabajo que contaban con su propio ordenador personal y una inmensa pantalla capaz de realizar buenos y rápidos diseños gráficos. Como disponían de buen material gráfico y mucha potencia para los ordenadores personales, se estableció una buena relación entre el usuario y el ordenador. Por ejemplo, todos los ordenadores incluían un dispositivo manual llamado *ratón*, que se podía deslizar por la superficie de un escritorio

como un disco de hockey y que percibía este movimiento a través de una bolita giratoria que sobresalía ligeramente de su base. El movimiento del ratón estaba coordinado con el movimiento de una flecha que aparecía en la pantalla del ordenador, y el usuario podía señalar a la máquina cualquier punto de la pantalla. Esto simplificó las cosas. En vez de escribir la opción deseada del menú, lo único que había que hacer era señalarla.

Cada una de las líneas del menú llevaba al lado una pequeña imagen gráfica, lo cual facilitaba muchísimo la tarea visual de elegir la opción adecuada y hacía que el sistema fuera parcialmente independiente del lenguaje. En las versiones posteriores, las imágenes, a las que se denominaba *iconos,* pasaron a ser la representación dominante y las palabras se relegaron a un plano inferior. Finalmente, los iconos se podían mover por toda la pantalla como objetos y colocar dentro de otros iconos que funcionaban como contenedores o marcadores para dispositivos tales como la impresora. Las interfaces de los iconos eran eficaces y fáciles de usar tanto para los principiantes como para los expertos, probablemente porque copiaban de los seres humanos la capacidad de manipulación de los objetos.

El proporcionar a cada usuario un ordenador independiente tenía muchas implicaciones, algunas de las cuales provocaban confusiones. ¿No representaba un paso atrás? Algunos de los pioneros del concepto de tiempo compartido eran de esa opinión. Los entusiastas del PARC alegaban que el tiempo compartido era muy imperfecto, si se pensaba en la comodidad del usuario. Inmovilizaba al usuario habitual, ya que lo limitaba a un terminal fijo conectado físicamente a una máquina inmensa. Y, lo que es peor, los sistemas de tiempo compartido nunca habían cumplido sus promesas iniciales. El elevado número de usuarios del ordenador y el alto costo del sistema en sí hacían que la gente a veces tuviera que esperar segundos e incluso minutos a que la máquina respondiera a las preguntas más triviales. Para las funciones de rutina, parecía más factible utilizar una pequeña máquina personal, especialmente teniendo en cuenta que el precio de los ordenadores bajaba rápidamente.

Alan Kay, el gurú del grupo PARC, inicialmente pensó en un ordenador personal de alta sensibilidad, del tamaño de un libro (apodado Dynabook por sus cualidades dinámicas), que tendría una pantalla en color de alta resolución y estaría unido por radio a una red informática mundial. El Dynabook sería mucho más que un simple ordenador, pues haría las funciones de secretaria, correo, biblioteca de referencia, centro de ocio y teléfono. Esta idea estaba y sigue estando más allá de las posibilidades de la tecnología, aunque esté escribiendo estas palabras con un ordenador portátil del tamaño de un libro, sentado en la recepción de un hotel.

Los costosos ordenadores personales del PARC tenían el tamaño de un escritorio, no el de un libro, hacían los gráficos en blanco y negro y ofrecían menos potencia de la que hubieran deseado sus diseñadores. Sin embargo, representaron un paso adelante hacia el concepto de Dynabook. Al principio se les llamó Interim Dynabooks pero, por suerte, posteriormente se les dio el nombre más musical de Alto.

Xerox tardó bastante en comercializar los descubrimientos del grupo PARC, aunque a finales de la década de los setenta creó un costoso centro de trabajo, llamado Star, que reflejaba muchas de las ideas del PARC. Después de diez años de estudios, algunos de los entusiastas del PARC se frustraron a causa de la lentitud empresarial de Xerox y tuvieron una favorable acogida por parte de Steve Jobs, que posteriormente sería cofundador de la empresa Apple. El resultado de dicho encuentro fue, algunos años después, el Apple Lisa y, aún más tarde, apareció su «hermanito» Macintosh, calificado como el ordenador personal «para todos». Aunque, por lo que se refiere a capacidad, sigue estando lejos del Dynabook, abrió un nuevo camino para que millones de personas entraran en el campo de la informática. E hizo posible que se abriera una nueva etapa en la relación entre seres humanos y ordenadores. A finales de los ochenta, casi todos los sistemas operativos nuevos para ordenadores «gráficos» se diseñan con el estilo Macintosh, de cara al usuario.



Gafas mágicas (modelo primitivo)
La instrumentación en el campo de la aviación militar —en la que el acceso instantáneo a la navegación, los sensores y los datos de las armas es asunto de vida o muerte— ha evolucionado y se ha convertido en las gafas mágicas. Este modelo se creó en 1986 para un proyecto de helicóptero experimental Boeing-Sikorsky. Los datos de todos los instrumentos del avión se proyectan en el campo visual del piloto. Por ejemplo, las señales luminosas del radar aparecen en la posición real donde se encuentran los objetos que se están rastreando.

GAFAS MÁGICAS

El interfaz gráfico que hace que el Macintosh y sus imitadores sean mucho más agradables de usar que las máquinas anteriores demuestra el valor que tiene hacer participar a las capacidades sensoriales de los seres humanos en el diálogo que se establece entre los hombres y sus máquinas. El Dynabook de Alan Kay era maravilloso en muchos sentidos, pero sería incapaz de ir mucho más allá, en esta dirección no verbal, que

los sistemas actuales, debido a sus limitaciones físicas, en especial el tamaño de la pantalla. Lo mismo que sucede con un libro normal, una de las características claves del Dynabook es que es portátil. Muchos de los usos que se sugirieron dejarían de existir si estuviera solamente en lugares fijos, como en el hogar y en el trabajo. ¿Se podrá conservar esta característica si, al mismo tiempo, se aumenta la participación sensorial del propietario? En otras palabras, ¿podemos imaginarnos un ordenador que presente la movilidad del ser humano y a la vez nos permita utilizar su memoria superior, su poder de cálculo y su amplia gama de comunicación?

Es evidente que sí. Sin embargo, no tendría el formato de un libro, sino el de un ajuar de alta tecnología. La prenda clave de este ajuar es un par de *gafas mágicas* o, en las primeras fases, gafas submarinas o un casco. Se apoyan en la nariz, como cualquier par de gafas, y llevan una impresionante variedad de instrumentos:

- Pantallas en color de alta resolución, una para cada ojo, con sistemas ópticos que cubren todo el campo de visión y hacen que la imagen del ordenador esté enfocada a una distancia adecuada. Las gafas *pueden* hacerse transparentes.
- Tres cámaras de televisión. Dos de ellas, de alta resolución, con lentes de gran angular, están situadas lo más cerca posible de los ojos, de tal manera que se puede ver hacia dónde se va cuando las cámaras están conectadas a las correspondientes pantallas de las lentes. La tercera, quizá de un muy gran angular, mira hacia atrás para registrar la cara del que las lleva puestas. Esto le permite al ordenador, y a cualquier persona que se ponga en comunicación por medio de un videoteléfono, controlar las expresiones faciales de quien las lleve puestas.
- Micrófonos y pequeños auriculares en la montura.
- Un sistema de navegación que registre continua y exactamente la posición y la orientación de las gafas (y, en consecuencia, de la cabeza del que las lleva puestas).

- Un potente ordenador capaz de generar imágenes, sonidos y voces sintéticas aunque realistas; de entender las órdenes habladas y de identificar y seguir objetos en el campo de visión de las cámaras.
- Un sistema de enlace de alta velocidad entre los bancos de datos de una red informática mundial, las bibliotecas electrónicas y las gafas mágicas que lleven otras personas.

No hace falta ser un experto para darse cuenta de que esta lista es demasiado amplia. Sin embargo, todas estas funciones existen y han sido objeto de muchas investigaciones. Reducir todavía más esta estructura es un reto posible en esta época de los Walkmans y de los televisores de bolsillo. Las gafas permiten que el ordenador controle lo que usted ve y oye según su petición verbal, y que siga sus movimientos. También le permiten observar el movimiento de sus manos y sus expresiones faciales, aunque no puedan controlar lo que usted percibe con las manos. Como el tacto es una fuente de información sobre el mundo, sería beneficioso que el ordenador nos permitiera conocer el tacto de las cosas, incluso de aquéllas que no están en nuestro entorno inmediato. Aparecen en escena los *guantes mágicos.*

Lo mismo que las gafas, los guantes son una demostración de la genialidad de la brujería tecnológica. Cada uno de los dedos del guante contiene una red de elementos que producen, sobre el dedo del que los lleva puestos, sensaciones de presión y de temperatura. Los experimentos psicológicos han demostrado que se pueden crear impresiones realistas. Por ejemplo, la presión uniforme y la sensación de frío se interpretan como «tengo los dedos sumergidos en agua». Los guantes tienen motores que actúan sobre las articulaciones de los dedos de tal manera que el que los lleva puestos puede percibir la resistencia a los movimientos. Estos mismos mecanismos le permiten al ordenador controlar la posición de los dedos.

Los guantes mágicos, lo mismo que las gafas, tienen sus limitaciones. Solamente pueden generar resistencia a los movimientos de los dedos, y la manipulación supone que todo el

brazo se pone en movimiento. Imaginemos un *abrigo con motor* que le proporcionara a las articulaciones del brazo la misma sensación de presencia que el guante le comunica a la mano. Si lo comparamos con los guantes y, sobre todo, con las gafas, el abrigo sería la imagen de la simplicidad. Pero en las primeras versiones del ajuar mágico, el abrigo puede ser un almacén conveniente para todo el *hardware* que no quepa ni en las gafas ni en los guantes, más pequeños y complejos. También puede estar unido a un punto inmóvil: un «sillón». Después, el abrigo se podrá trasladar con la misma facilidad que si fuera un traje espacial y, finalmente, parecerá un traje de gala.

Lo mismo que sucede con los ordenadores actuales, el maravilloso *hardware* que hemos esbozado es tan útil como su *software*. En las secciones siguientes le presentaremos algunos programas que posiblemente le apetecerá probar cuando lleve su nuevo y elegante atuendo.

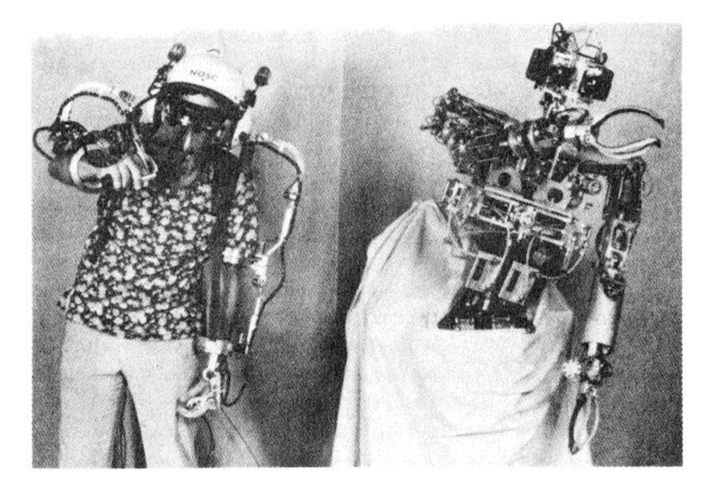
ENCONTRARSE A UNO MISMO

No sé si a usted le pasará lo mismo, pero yo me pierdo con frecuencia. Como carezco del sentido de la orientación, suspiro por tener en el bolsillo algo que me ayude, que no sólo me indique dónde estoy (los mapas electrónicos que hay ahora en los coches ya lo hacen) sino que me guíe hasta mi punto de destino, sepa dónde he estado y se acuerde de la lista de la compra cuando note que estoy cerca de la tienda. Las gafas tienen un sistema de navegación, probablemente una combinación de un mecanismo que mide la distancia al suelo y hace girar unas balizas por radio y otro que percibe las fuerzas de aceleración para deducir el movimiento. Lo haga como lo haga, lo cierto es que este mecanismo sabe en todo momento dónde está usted, que lleva las gafas puestas.

Robin y su familia han cambiado de domicilio y ahora viven en una región muy apartada. Este territorio es desconocido y, por una vez, usted empieza a sentir aprecio por esa brillante línea verde que ha creado el programa Carretera de

Adoquines Amarillos y que le guía por la autopista. La línea se sitúa en el carril de la derecha y se dirige hacia una desviación situada a unos ochocientos metros. Por si acaso no se ha dado cuenta, aparece a la derecha una flecha que lanza destellos y flota sobre la desviación, mientras el programa le susurra al oído: «La desviación se aproxima, reduce a cincuenta.» Un par de kilómetros después, se acaba el asfalto y el piso es de grava. Para prevenirle, la guía cambia de color y se pone amarilla, color que indica más prudencia. Luego, la línea se vuelve roja, y la carretera se convierte en un lodazal. El programa anuncia «Pasar a un mapa particular», lo que indica que en la base de datos de Rand McNally no aparece este sendero. Los datos provienen de los archivos personales de Robin. «Quedan seis kilómetros para llegar a casa de Robin», anuncia el programa, que interrumpe, momentáneamente, una antigua canción de los Beatles.

Está oscureciendo, «Empezar a telefonear —Robin», le dice al ordenador que tiene sobre la nariz. Se pone en funcionamiento el programa telefónico, que conecta por medio de la red celular o de satélite. El programa Carretera de Adoquines Amarillos sigue adelante por su cuenta. «Estoy llegando», le dice el visitante a Robin. «Fenomenal. La cena ya casi está. ¿Prefieres arroz o patatas?» «¿Pueden ser las dos cosas?» «¡Claro! ¡Tú verás! Hasta ahora.» «Hasta ahora mismo.» El programa telefónico termina. La guía roja oscila por entre los árboles. Las luces están puestas, pero no se ve bien. De repente, el coche da un bandazo a la izquierda. Ha chocado contra algo, a lo mejor con una piedra que había en la carretera, y la rueda delantera se ha quedado atascada en un bache lleno de barro. Al intentar salir, se hunde más. «Saca el mapa», se pide al programa de navegación. Y el mapa pedido aparece en el aire delante de usted. La imagen de su coche en el mapa le indica que está a unos tres kilómetros de su punto de destino, ya que la carretera tiene muchas curvas; pero está a menos de medio kilómetro si va por un sendero directo que atraviesa el bosque.



Un agente robot
Este agente robot se creó en 1986, en el Centro de Sistemas Oceanográficos y Navales de Hawái. Los movimientos que hace el que lo maneja, a la izquierda, los copia el robot de la derecha, y las imágenes que captan los ojos del robot, que son cámaras de televisión, se envían a las (enormes) gafas mágicas que lleva puestas el ser humano, el cual tiene la sensación objetiva de encontrarse en el cuerpo del robot.

«Hola, Robin, soy yo otra vez. He tenido un pequeño accidente.» Después de pensarlo, Robin le sugiere que sea un poco aventurero. «Nos podemos ocupar del coche por la mañana. La semana pasada, Marty y yo descubrimos un atajo y estás muy cerca. Si vienes andando, no tardarás más de media hora.» Como usted es valeroso, acepta. El disco del sendero pasa del ordenador de Robin al suyo a través de la conexión telefónica. Su programa de navegación une la ruta que va por el atajo con los datos de la carretera, y a la guía original la sustituye otra que va un trecho a lo largo de la carretera y luego se interna en el bosque. Cierra el coche con llave y sigue la línea.

Está muy oscuro, así que activa su programa de visión nocturna: las cámaras de los ojos funcionan al máximo de sensibilidad y le proporcionan mejores imágenes a las pantallas que hay a la salida del sistema de control de la navegación. Observa que la guía roja va por una senda que atraviesa los bosques. No sucede nada, excepto que se hace algunos ligeros arañazos.

La cena estaba caliente. Y deliciosa.

SITIOS A LOS QUE SE PUEDE IR

El transporte y la comunicación han mejorado notablemente en los últimos quinientos años. Sin embargo, la geografía, aunque ha dejado de ser el factor fundamental que condiciona el comercio, sigue poniendo límites a cómo y con quién haremos ciertas cosas. Las diferencias que existen entre el transporte y la comunicación irán desapareciendo a medida que vayamos siendo capaces de proyectar nuestra capacidad y nuestro conocimiento para reconocer lugares. Cuanto más sencilla sea esta proyección, mayores serán las probabilidades de que los intereses comunes sean la base fundamental de la relación entre ambas. El ajuar mágico se puede utilizar para visitar lugares reales del mundo y lugares «irreales» emplazados en alguna de las simulaciones del ordenador.

La forma más evidente de presencia remota supone la existencia de un robot físico, un robot lejano que usted pueda controlar por medio de la red global de comunicaciones. Las gafas mágicas le permitirán ver por los ojos del robot, el abrigo y los guantes le permitirán sentir, gesticular y actuar por medio de los manipuladores del robot, y los controles de los pies que tiene en su sillón le permitirán desplazar al robot. Si alquila «agentes» robots en localidades muy remotas, podrá visitar, hablar y colaborar en proyectos que estén diseminados por todo el mundo sin tener que abandonar las comodidades de su hogar y sin tener que arriesgar su integridad física o exponerse a viajes largos y tediosos.



Propiedades inmobiliarias irreales.La carretera a Punto Reyes
Esta escena se sintetizó en el departamento de diseño gráfico por ordenador de Lucasfilm en 1985, con un modelo de ordenador básico tridimensional. Con unas gafas mágicas y la potencia informática adecuada podremos deambular por mundos de fantasía como éste. © 1986 Pixar.

Lo que más se parece a la idea del agente es un agente *humano.* Alguien que se desplazara a un sitio interesante y llevara una especie de equipaje mágico con el que pudiera transmitir al sillón viajero lo que viese, escuchase, tocase e incluso oliese. En su forma más simple, la relación es unidireccional, del agente al observador pasivo. Este tipo de comunicación unidireccional se puede grabar y volver a repetir a voluntad, lo que nos proporcionaría una forma verosímil de la extensión «sensorial» de la «película». Pero la conexión también puede ser bidireccional, con movimientos e impresiones sensoriales —bien editada por un programa inteligente—, y transmitirse en

ambas direcciones. Lo más probable es que los participantes sin experiencia se encuentren tan faltos de coordinación, y se diviertan tanto, como en una carrera con tres piernas, pero, con la práctica, se podrá hacer un buen trabajo de equipo y los consejos del experto llegarán al lugar donde se ha planteado el problema. El «agente de campo» controlará los movimientos la mayor parte del tiempo, y el participante del sillón observará, escuchará, percibirá y aconsejará. Pero cuando la tarea exija una habilidad manual que domina mejor el que está en casa, el agente de campo se relajará y permitirá que sus motores, controlados a distancia, hagan el trabajo, como si estuviera poseído por un espíritu.

Las reuniones con los agentes no tienen que celebrarse necesariamente en el mundo real, ya que con la simulación del ordenador se pueden hacer cosas mejores. Las «propiedades inmobiliarias irreales» que crea el ordenador no tienen límites intrínsecos, ni en extensión ni en características físicas. Es un mundo distinto, en el que la magia es una rutina. Y las pantallas de los ordenadores de hoy nos permiten asomarnos a él: los diseñadores de coches examinan los futuros modelos, los físicos contemplan el interior de las explosiones nucleares, y los usuarios de Macintosh reorganizan sus ficheros en un espacio irreal llamado «la superficie del escritorio». Si tenemos un ajuar mágico, podemos penetrar en estos mundos y explorarlos desde dentro.

La modesta casa de sus sueños se está convirtiendo en realidad. Hace algunas semanas estuvo hablando con el arquitecto, y hoy recibe la llamada que esperaba. «¡Hola! ¿Qué tal? Ya está hecho el proyecto. ¿Tiene tiempo para echarle un vistazo?» «¡Claro que sí!» «Bueno, cambiamos de escenario. Ahora es la Escena M5.» Después de algunos segundos, el ordenador de sus gafas pregunta: «¿Permites que cambie el panorama?» «Te lo permito», responde usted.

Le rodea una lograda reproducción de la parcela y sus alrededores. «Creía que habíamos puesto la casa aquí. Así queda espacio para dos trozos de jardín, uno delante de la casa

y otro detrás, y el camino y el garaje caben a la derecha.» Aparece en el suelo un esbozo de la casa. «En el primer piso están el vestíbulo, un aseo para los invitados, el salón, el comedor, la cocina y el hueco de la escalera.» En el esbozo aparece un plano del piso. Le da la impresión de que la idea del arquitecto es más complicada que la de usted y que incluye pantallas de control. «Vamos a levantar el primer piso.» Del plano del piso surgen las paredes. Ahora usted ve un segundo plano del piso sobre el edificio truncado. «En este nivel irán dos alcobas, dos cuartos de baño y un despacho.» El segundo piso se pone en su sitio. «El tercer piso será el ático, y habrá posibilidades de poner dos alcobas y un cuarto de baño.» Se coloca el tejado y se añade un poco de paisaje. Usted vuelve al piso de abajo y se pasea por la casa. «¿Podríamos hacerla de ladrillo?» Los muros de piedra pasan a ser de ladrillo visto. «El jardín de atrás me parece un poco pequeño. ¿Se puede mover todo hacia delante?» «No podemos salimos mucho de la línea, pero te puedo dar unos cinco metros.» La casa se desliza hacia delante. «Así está mejor. Vamos dentro.»

Se abre la puerta principal. Usted observa que hay un interruptor a la derecha y desea encender la luz. Una figura, caricatura de una mano, lo hace y el vestíbulo queda iluminado. La vista desde las ventanas del salón no es precisamente inspiradora. «¿Podría haber aquí una ventana salediza en vez de ésta?» «Un momento, tengo que construirla.» Se da una vuelta por la cocina, que parece bastante amplia, y por el comedor. «Ya está la ventana.» El salón está mucho mejor con la luz del sol que entra por la ventana salediza. «¿En qué estación estamos?» «En la actual. Vamos a darnos un paseo por todo el año.» La iluminación cambia, de la mañana a la tarde, recorriendo las cuatro estaciones. En invierno es especialmente acogedora.

Después de un paseíto por el piso de arriba y de algunas modificaciones en el color de las paredes, se despiden. «Le dejaré una copia. Podrá hacer cambios pequeños y colocar los muebles. No se preocupe por el diseño, no lo puede estropear.

La máquina no le dejará que haga ninguna tontería.» «Mi familia se va a quedar encantada. Hablaremos después.»

PROGRAMAR CON LAS MANOS

Los expertos en distintos campos aseguran que *ven* y *sienten* el tema de su trabajo cuando piensan en él. Esto no tiene mucho mérito en trabajos relacionados con situaciones y objetos físicos como la escultura, el deporte o el diseño mecánico, por ejemplo. Es más sorprendente en campos supuestamente abstractos como la música, el lenguaje, las matemáticas o la física teórica. Sin embargo, Einstein afirmaba que podía sentir en los brazos y en el cuerpo el significado de sus ecuaciones como si se tratara de objetos sólidos.

Como sugerí en el Capítulo I, parece que las zonas sensoriales y motoras del cerebro humano, grandes y muy evolucionadas, son la central eléctrica oculta que está tras el pensamiento humano. Como estas estructuras, que tienen mil millones de años de antigüedad, son altamente eficaces, pueden expresar un millón de veces la potencia informática efectiva de la parte consciente de nuestra mente. Pero mientras que se puede conseguir el rendimiento del principiante empleando solamente el pensamiento consciente, para llegar al nivel de maestro hay que recurrir a los enormes recursos que ocultan esas zonas antiguas y especializadas. Algunas veces, parte de esta potencia se puede utilizar para encontrar y desarrollar un camino entre un problema y una intuición sensorial.

Aunque algunos individuos, gracias a afortunadas combinaciones de herencia y oportunidad, tienen intuiciones certeras en algunos campos, la mayor parte de nosotros no somos más que aficionados. Lo que necesitamos para mejorar nuestro rendimiento es que haya metáforas externas explícitas que hagan aflorar nuestras capacidades instintivas de una forma directa y repetible. Para aumentar nuestra comprensión y nuestra retentiva utilizamos gráficos, reglas, modelos físicos que reproducen relaciones y otros mecanismos. En fechas más recientes,

las interfaces interactivas gráficas de los ordenadores, como las de Macintosh, han acelerado mucho el aprendizaje de los principiantes y hacen que el manejo de la máquina les resulte más sencillo a los que tienen experiencia. La relación sensorial completa, que es posible con las gafas mágicas, nos puede permitir que vayamos mucho más lejos por este camino. Acaso el encontrar las mejores metáforas será el trabajo de toda una generación. De momento, nos podemos divertir haciendo conjeturas.

Tiene ante usted el paisaje familiar de la parte superior de su sistema de archivo. En primer plano, sobre un prado de hierba verde, se pueden ver algunas rocas de distintos tamaños, colores y formas que llevan etiquetas: «Presupuesto», «Dibujos», «Juegos» y otras. En la distancia, cubiertas por la neblina, se perciben unas grandes colinas engalanadas con inscripciones de colores: «Oxford English Dictionary» y «Enciclopedia Británica». Si se dan dos golpecitos sobre la roca que lleva la etiqueta de «Espacio», se hace más grande y se abre una puerta en uno de sus lados. A través de la puerta se ve lo que podría ser un cinturón de asteroides. Una de las rocas que flota en la oscuridad va marcada con la etiqueta «Curvas del cielo». Se acerca a ella, llama dos veces y entra. Contempla una tierra hermosa, blanca y azul, y algunos escombros de distintas formas y más feos. Es un proyecto inacabado, y algunos de los experimentos que ha realizado y no han tenido éxito precisan todavía de cierto tiempo de reposo.

El problema de hoy consiste en crear una simulación para un cable, largo y fuerte, que está en órbita alrededor de la Tierra. El cable tiene masa y una cierta elasticidad. Usted ha aprendido que se puede aproximar uniendo una gran cantidad de resortes simples y pesos todavía más simples. Un resorte simple une dos puntos y ejerce sobre ellos una fuerza proporcional a la elongación a partir de una cierta longitud. La posición y la velocidad de una masa varían en proporción con la fuerza que se les aplica, de acuerdo con la Tercera Ley de Newton. La fórmula, para un resorte, es $F = K (I - I_0)$*, donde*

I es la longitud actual del resorte e I_0 es la longitud inicial, antes de que se estire. K es la constante del resorte: cuanto mayor es K, más difícil es estirar el resorte. F es la fuerza que ejerce el resorte en sus extremos. Esta relación es la que existe entre los escombros que estropean el paisaje. Usted comienza eligiendo algunos de los componentes que ha construido en el pasado. Un punto es como una motita negra que flota en el espacio (aunque haya tres números que son los que le asignan sus coordenadas X, Y y Z. Pero eso es de la lección de ayer). Si da un golpecito en el punto y dice «¡Duplicar!», aparece otro. Los dos puntos representarán los dos extremos del resorte.

Usted coge una larga flecha. Es como una línea muy larga con una punta de flecha en cada uno de sus extremos y un número (su longitud) en el centro. Coloca las dos puntas sobre los extremos y calcula la distancia que existe entre ellos. Ahora coge uno de los puntos y lo mueve. La flecha sigue al punto y el número que indica la longitud cambia, obediente. Dándole un golpecito, usted dice: «Llamo a esto I.» Y la dimensión cambia al símbolo I. «Vete a la fórmula del resorte», ordena usted. Y aparece por el paisaje una copia de la fórmula, se coloca cerca y empieza a responder a la distancia entre los puntos. Aparecen ahora los huecos para Kel0y usted les da valores. Otro de los componentes ya preparados es «Vectorizar». Dados un par de puntos y una magnitud simple, le asigna dirección a la cantidad, es decir, la dirección de una línea que una los dos puntos. La orden de «Vectorizar» está unida a los dos puntos y se ejerce en direcciones opuestas. Usted ordena ahora: «Vectorizar F.» Los puntos ejercen la fuerza del resorte, aunque permanecen fijos. Cuando usted agarra uno u otro, da tirones dentro de su mano. Cuanto más lejos se lo lleva, más fuerte tira en dirección al otro punto, que sigue fijo. Este tipo de método experimental sólo se puede utilizar con algunas magnitudes. Otras serían la posición, el color y la temperatura. En muchas aplicaciones, resulta muy útil pasar de medidas más abstractas a otras más palpables.

Ahora añade masa a los extremos. Esto les permite moverse con independencia, bajo el control del momento y de las fuerzas aplicadas como, por ejemplo, la del resorte. Si los extremos están libres, el resorte vibra. La vibración no disminuirá hasta que usted no añada un término de amortiguamiento a la ecuación de la fuerza; este término dependerá de la velocidad de cambio de I. Ahora el resorte se comporta de manera razonable y usted lo estira varias veces para divertirse. Si invoca al compilador, el resorte se convierte en un simple objeto y aumenta mucho la eficacia del programa. Ahora edita la imagen del resorte de forma que parezca una espiral alargada que tiene en los extremos unos discos negros que representan la masa. Después se unen, extremo con extremo, una docena de resortes para formar una especie de cuerda. Su simulación se encuentra en un buen punto, pero ya es hora de comer. Después del almuerzo, pasará a otra sección, cambiará los parámetros de las distintas partes y a lo mejor representa en un gráfico los alargamientos de las distintas secciones. Después aplicará el modelo gravitatorio terrestre, pondrá la cuerda en órbita y observará lo que sucede.

FÍSICA ELEMENTAL

Sócrates, cuyas enseñanzas han llegado hasta nosotros gracias a Platón, que las recopiló, nunca escribió ningún libro. Debía pensar que no estaba bien escribir, ya que uno puede hacer alarde de tener amplios conocimientos y no saber nada en realidad. La misma capacidad de recordar y de pensar sobre los recuerdos era arriesgada. Además, un argumento que se presenta en un libro no se puede rebatir si no se está de acuerdo con él, mientras que con una persona se puede discutir o nos puede aclarar sus razonamientos. Las dos objeciones son razonables. El conocimiento que se obtiene de los libros es, evidentemente, algo seco y estático si se compara con el conocimiento activo en la mente de una persona inteligente. La invención de la imprenta contribuyó a agravar la situación. Y, sin embargo, los

libros tienen un alcance, una capacidad y una permanencia mucho mayor que la de la memoria de cualquier persona, propiedades éstas que han hecho posible la civilización moderna.

Antes de que se inventara la imprenta, los libros eran artículos caros, que se hacían trabajosamente a mano y que sólo se podían encontrar en las pocas y dispersas bibliotecas que existían. Era muy difícil conseguir ejemplares para uso particular, y los eruditos se tenían que aprender de memoria volúmenes enteros. Las ayudas artificiales para la memoria se consideraban como algo muy valioso, y era corriente que los rivales teológicos, políticos y comerciales las guardaran celosamente. Una de estas técnicas, muy eficaz, y que tuvo innumerables variaciones a lo largo de toda la Edad Media, era El Paseo. Se recordaba o imaginaba un edificio muy grande, por ejemplo, una catedral con muchas capillas. Entonces se recitaba el libro o lectura que había que aprender de memoria mientras que, al mismo tiempo, se paseaba mentalmente por la estructura elegida. Cada una de las capillas de la catedral, o una zona de las mismas, servía a los ojos de la mente como recordatorio de una parte del texto, acaso porque hubiera en ella algún objeto que recordara a uno de los temas tratados en la lectura. La tarea se dividía en porciones. Y recordar las zonas, aisladas unas de otras, no requería un gran esfuerzo memorístico. Se podía reconstruir mentalmente el edificio entero volviendo a pasear por él, visitando las salas una por una. Y las imágenes mentales que se generaban hacían recordar las partes de texto asociadas.

El Paseo puede ser muy efectivo porque enmarca la nueva necesidad cultural de aprender de memoria grandes cantidades de palabras en la capacidad de recordar, mucho más antigua, donde vimos o dejamos distintas cosas. El recordar dónde se encuentra una fuente de alimentos, un refugio, un peligro, un amigo, un enemigo o, simplemente, un lugar conocido tiene muchas ventajas en la vida cotidiana y muchos de nosotros lo hacemos con naturalidad y bastante bien. Al menos una parte

de nuestra memoria está organizada de una forma que podríamos llamar geográfica, para facilitar un tipo de recuerdo en concreto.

Una lección que nos llegue por medio de una versión avanzada del ajuar mágico puede ser, a la vez, tan interesante como un diálogo personal y tan permanente como un libro impreso, y adaptarse mucho más a las capacidades naturales que cualquier otro método que existiera anteriormente.

Se abre el pórtico sobre el que se lee «Gravedad», y deja ver una escena pastoril brillantemente iluminada por la luz del sol Una carretera rural sombreada por dos filas de árboles serpentea por las lejanas colinas, salpican el cielo nubes algodonosas y se escucha el piar de los pajarillos. Algunos árboles son manzanos y, de vez en cuando, una de sus frutas cae al suelo. En la distancia, por la carretera, se percibe la figura de un ser humano, con peluca, que está sentado bajo uno de los manzanos. Usted reconoce a sir Isaac Newton. Está igual que en el retrato del capítulo dedicado a las «Leyes del Movimiento».

«Hola, joven», dice sir Isaac. «He estado devanándome los sesos pensando en la naturaleza de la atracción entre la Tierra y diversos objetos. Esta manzana, por ejemplo, tira de la mano con una cierta fuerza.» Le da la manzana y usted comprueba que realmente pesa. «Una manzana que tenga el doble de sustancia que ésta empuja la mano con el doble de fuerza.» La manzana se hace más grande y pesada. «El gran Galileo observó que, cuando se deja libre, un objeto cae hacia el suelo y su velocidad aumenta deforma constante, con independencia de su peso.» Usted recuerda las demostraciones de Galileo con las bolas que caían en las «Leyes del Movimiento». «Sí, sí, continúe.» Newton, con el ceño ligeramente fruncido, prosigue. «Podemos deducir que cada una de las partículas de un objeto se siente atraída hacia el centro de la Tierra con una fuerza que es proporcional a su masa. ¿Se modifica esta atracción con la distancia a la Tierra? Se puede suponer que la influencia se extiende a grandes distancias y que alcanza a la Luna

en su viaje mensual. Si se pueden aplicar las mismas leyes a los cuerpos celestes y a los de la Tierra, entonces nuestros estudios del movimiento de los objetos indican que una fuerza ejercida en la dirección del centro de la Tierra basta para curvar la trayectoria de la Luna. Y', sin embargo, la fuerza necesaria es casi cuatro mil veces menor, por partícula de masa, que la que experimenta la manzana que usted sostiene en la mano.»

A medida que habla, el suelo se va inflando a gran velocidad y usted, Newton y el árbol se encuentran en la cima de una colina que se eleva como un cohete. «Consideremos la trayectoria de un objeto que se arroja horizontalmente desde una gran altura. Su manzana, por ejemplo.» Usted acepta la sugerencia y arroja la manzana levantando graciosamente la mano. (En realidad, los motores de su chaqueta y de sus guantes zumban momentáneamente mientras se oponen al movimiento de su brazo, simulando las fuerzas de inercia de la manzana.) La manzana describe un arco hacia el suelo y cae cerca del horizonte. La colina ha dejado de crecer, pero usted se encuentra muy arriba y puede apreciar la forma esférica del planeta. Puede inventar diversos trazados de los continentes. Evidentemente, es una reproducción a escala y en miniatura de la Tierra. Sir Isaac le da otra manzana y le recomienda que la arroje con más fuerza. Su arco se pierde más allá del horizonte en una curva que va casi paralela al suelo. Usted escucha, a través del suelo, el ruido que hace la manzana al despedazarse bajo sus pies. Y ahora, con más fuerza todavía, arroja otra manzana, pero esta vez no hay impacto y al cabo de un rato la manzana pasa zumbando por detrás de su cabeza y recorre su camino una vez más. Una vista lateral en miniatura de usted, la colina, la Tierra y las manzanas facilita su comprensión. Cada trayectoria describe una elipse que regresa a su punto de partida si no choca primero con el suelo. Newton le recuerda las leyes de Kepler sobre el movimiento de los planetas y afirma que, con respecto a las manzanas, sólo se cumplen si la atracción decrece con el cuadrado de la distancia desde el centro del planeta. Usted es escéptico; así que ambos deciden

experimentar con otras leyes. En algunos lanzamientos, las manzanas describen trayectorias que no se vuelven a repetir. En otros sí describen una elipse, pero no cumplen la segunda ley de Kepler', que dice que la línea que une el centro del planeta con el objeto en órbita tiene que barrer áreas iguales en tiempos iguales. Al final ya se le está cansando el brazo, y decide comunicarle que le ha convencido.

Pero a veces su propio escepticismo da lugar a preguntas que dejan perplejo a su anfitrión. Recuerda que una vez Newton le respondió: «Un interesante rompecabezas. Déjeme que lo piense.» Y un día se presenta, sin aliento, con los faldones de la casaca flotando al viento, sujetándose la peluca con una mano y seguido por una nube de polvo. (Usted se imagina que el software *del libro no estaba preparado para contestar a esta pregunta la primera vez que la planteó y que ha enviado un mensaje a los autores del libro. Los autores han hecho algunas entradas en la base de datos del libro para que Isaac pueda contestar a su pregunta del otro día y a otras semejantes que le puedan hacer en el futuro.)*

La colina se encoge y vuelve a quedarse plana, y usted se encuentra de nuevo en la carretera. La siguiente vez que se detiene es en un prado, en el que se discutirán las partes más formales de la lección. Ya se han reunido allí una media docena de exóticas criaturas. Hay mucha gente de todo el mundo que está leyendo este libro y la red mundial hace que sea posible que cada persona sea consciente de la presencia de las demás mientras se está realizando el estudio. En este tipo de asociaciones, la mayor parte de la gente se aprovecha de la libertad que concede la simulación para que sus cuerpos asuman formas diferentes con el objeto de conservar el anonimato o por capricho. En su grupo hay un Lobo, un Ojo Flotante, un Hombre Flaco, una Mariposa Gigante, un Dragón y un Tanque Pequeño. Usted, sin ir más lejos, ha tomado la apariencia de un Enano, con bastante estilo, que lleva un hacha y un gorro con una borla. Algunos de sus compañeros, que empezaron este libro con usted, ya no están en su clase: o le han adelantado, o se han quedado atrás o se han separado en alguno de

los caminos laterales con los que se han encontrado. De vez en cuando, usted elige nuevos compañeros a medida que van apareciendo nuevas subclasificaciones. Es como si el mundo entero estuviese dividido en «pueblos de interés común» de este tipo. El grupo puede estar compuesto de dos personas o de miles. En algunas ocasiones, es muy agradable recorrer en solitario los caminos del aprendizaje. Entre otras ventajas, se puede organizar el trabajo como quiera, ya que no existen compromisos.

Después de la lección echa la mirada hacia atrás, por la carretera. A lo lejos se ve una estación de ferrocarril y hay parado un tren de pasajeros. Si se fija bien, se dará cuenta de que por una de las ventanillas se percibe, un tanto borroso, el perfil del científico más famoso del mundo. Pero se encuentra cansado y decide desconectar por hoy. La relatividad puede esperar hasta mañana.

IV. MIRANDO AL FUTURO

¿Qué es lo que sucederá cuando se pueda sustituir al ser humano por máquinas cada vez más baratas? O, dicho de otro modo, ¿qué es lo que voy a hacer yo cuando exista una máquina que pueda escribir este libro o hacer mis investigaciones mejor que yo? Estas preguntas revisten gran importancia para mucha gente de todas las profesiones y, dentro de algunas décadas, afectarán a todo el mundo. A causa de su diseño, las máquinas son ahora nuestras esclavas obedientes y habilidosas. Pero las máquinas inteligentes, por muy benévolas que sean, amenazan nuestra existencia, ya que pueden sustituirnos en el espacio ecológico que habitamos. Las máquinas que sean simplemente tan listas como los seres humanos tendrán enormes ventajas sobre nosotros en situaciones de competitividad. Su fabricación y su mantenimiento cuesta menos, así que se puede aumentar su número utilizando solamente los recursos de que disponemos. Se pueden mejorar notablemente y programarlas para que trabajen sin descanso.

Por si todas las creaciones tecnológicas no fueran bastante amenazadoras, la rapidez con que se producen las innovaciones representa un reto mucho más serio a nuestra seguridad. Nosotros evolucionamos a un ritmo muy lento, y pasan millones de años antes de que se produzcan cambios significativos. Las máquinas necesitan solamente décadas. Cuando haya multitudes de máquinas muy económicas que hagan el trabajo de los programadores y de los ingenieros, cuya tarea sea mejorar el *software* y el *hardware* que las hacen ser lo que son, aumentará todavía más la rapidez de las innovaciones. Y las sucesivas generaciones de máquinas así creadas serán todavía más inteligentes y más económicas. No existe ninguna razón para pensar que el nivel humano represente ningún tipo de tope superior.

Cuando las calculadoras de bolsillo sean más listas que los seres humanos, ¿qué es lo que será capaz de hacer un ordenador grande? Simplemente, nos habremos quedado fuera de juego.

Por lo tanto, ¿por qué meternos precipitadamente en una era de máquinas inteligentes? Creo que la respuesta es que no podemos hacer otra cosa, si pretendemos que nuestra cultura siga siendo viable. Las sociedades y las economías están tan sometidas a las presiones de la evolución competitiva como los organismos biológicos. Antes o después, los que tengan capacidad de expandirse y diversificarse con más rapidez, serán los que dominen. Las culturas compiten unas con otras por los recursos del universo al que tenemos acceso. Si la automatización es más eficaz que el trabajo manual, entonces las sociedades automatizadas serán más ricas y tendrán más posibilidades de sobrevivir en los momentos difíciles y de expandirse en los favorables. Si Estados Unidos detuviera unilateralmente el desarrollo tecnológico (una idea que se pone de moda de vez en cuando), sucumbiría en seguida ante el poder militar de naciones enemigas o ante el éxito económico de sus socios comerciales. En cualquiera de los dos casos, los ideales sociales que llevaran a esta decisión carecerían de importancia a escala mundial.

Si a causa de algún pacto bastante improbable el género humano en su conjunto decidiera renunciar al progreso, el resultado a largo plazo sería la extinción casi segura. El universo es un suceso al azar tras otro. Antes o después evolucionaría algún virus mortal de necesidad para los seres humanos, un gran asteroide chocaría con la Tierra, el Sol se expandiría, sufriríamos una invasión de las estrellas o un agujero negro se tragaría la galaxia. Cuanto más amplia es una cultura, más diversa y competente es, y más posibilidades tiene de detectar y enfrentarse con los peligros externos. Los sucesos más importantes ocurren con menos frecuencia. Si una cultura crece con la suficiente rapidez, tiene una oportunidad finita de sobrevivir para siempre. En el Capítulo VI divagaré sobre los esquemas que le per-

mitirían a un ente reestructurarse a sí mismo con objeto de funcionar indefinidamente, incluso en el caso de que su universo se extinguiera.

Antes de que pase mucho tiempo, el género humano se expandirá por el sistema solar y las colonias ocupadas por los seres humanos serán parte de esta expansión. Pero solamente podremos sobrevivir en las superficies de otros planetas o en el espacio exterior si desplegamos una cantidad enorme de maquinaria. Por ejemplo, con el proyecto Apolo algunas personas estuvieron en la Luna durante varias semanas y costó cuarenta mil millones de dólares, mientras que los módulos Viking funcionaron en Marte durante años y costaron solamente mil millones de dólares. Estas máquinas, tan capaces como el ser humano, que se construyeron para el proyecto Viking, han podido recopilar muchísima más información sobre Marte que las personas que estuvieron trabajando en la Luna, y la razón es que las máquinas se pueden diseñar para que funcionen cómoda y económicamente en medios distintos a la Tierra.

El espacio exterior es ya un campo muy lucrativo para los propietarios de satélites de comunicación. Cuando los gastos de transporte empiecen a disminuir, comenzarán a producir beneficios otras actividades. Las fábricas espaciales que procesen materias primas compradas en la Tierra o en las colonias espaciales, funcionarán con máquinas supervisadas por los seres humanos y la venta de esos productos dejará beneficios. Los elevados gastos que supone mantener a las personas en el espacio nos hacen pensar que en las colonias espaciales habrá más máquinas por ser humano que en la Tierra. Y a medida que las máquinas vayan siendo más capaces, la economía hará que la proporción entre máquinas y personas sea más elevada. En esta fase, los seres humanos no tendrán que ser forzosamente menos en número. Sencillamente, las máquinas se multiplicarán con mucha mayor rapidez y a cada nueva generación irán siendo más competentes. Imaginemos lo tremendamente rentables que serán las fábricas de robots que se podrán construir en los asteroides. Las máquinas funcionarán con energía solar, buscarán materias primas y las llevarán a inmensas plantas de

procesamiento, automáticas y abiertas. Los robots transformarán los metales, semiconductores y plásticos en componentes que servirán para montar otros robots o serán piezas estructurales de otras fábricas. Las máquinas se podrán ir reciclando cuando se estropeen. Y si la velocidad a la que se reproducen es mayor que la velocidad a la que se estropean, las fábricas crecerán exponencialmente, como una colonia de bacterias, en una escala brobdingnagiana. Y los inversores se podrán hacer inmensamente ricos con una pequeña fracción del volumen de producción de materiales, componentes y robots.

Finalmente, cuando se utilicen los descubrimientos científicos y técnicos de autorreproducción de mecanismos superinteligentes para hacer que las máquinas sean todavía más listas, los seres humanos, trabajadores, ingenieros de diseño, gerentes o inversores, llegarán a ser innecesarios en las empresas espaciales.

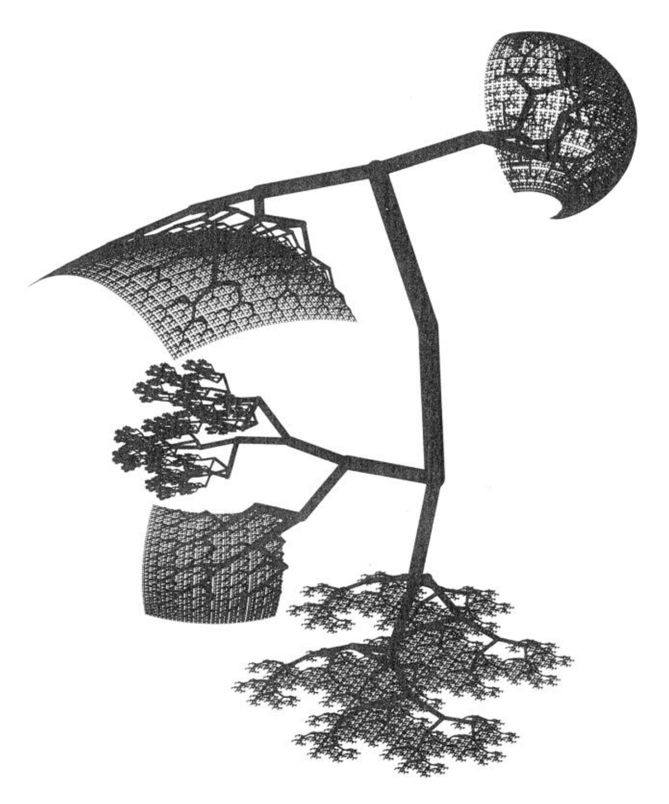
ROBOTS ARBUSTO

El mundo de los seres humanos lo han modelado manos humanas, que siguen siendo nuestras mejores y más efectivas herramientas. Y, sin embargo, hay muchas cosas útiles y que se pueden describir fácilmente que están más allá de nuestras posibilidades: tirar con fuerza de una cuerda por ambos extremos al tiempo que se aprieta el nudo entre los dedos, levantar el paquete y pasar la cuerda alrededor del mismo, muy tensa, cuatro veces. Si se intenta hacer esto, el grado de éxito es variable y hay que utilizar herramientas especiales.

No es probable que nuestros descendientes superinteligentes se conformen con los simples y cortos dedos. Tengamos en cuenta las siguientes consideraciones. Los gusanos y otros animales de forma esférica o alargada son incapaces tanto de manipular objetos como de desplazarse bien. Los animales con patas (un palo con palitos móviles) se desplazan bastante bien, pero siguen siendo torpes en el campo de la manipulación. Los animales como nosotros, que tienen dedos en las extremidades

(palos con palos que tienen palos) pueden manipular mucho mejor. Generalicemos ahora el concepto: un robot que tenga aspecto de árbol, con una raíz grande que se vaya dividiendo en ramitas cada vez más delgadas, cortas y numerosas, para terminar en un número astronómico de cilios. Todas las ramas intermedias pueden balancearse hacia delante y hacia atrás, y de derecha a izquierda, y la parte de arriba, de donde salen las siguientes ramas, gira sobre el eje de la rama. A lo mejor la rama puede variar de longitud, como si fuera un telescopio, y el número de movimientos de cada rama se puede cambiar para más niveles. Las uniones llevarán sensores para medir su posición y la fuerza que ejerce. Aunque la estructura global se componga de ramas, rígidas y mecánicas, tendrá una flexibilidad «orgánica», ya que todos sus componentes pueden moverse.

Un robot con este diseño se puede construir a sí mismo. Para que el proceso comience se pueden «sembrar» pequeños arbustos, cuyo peso sea de algunas millonésimas del que tendrá el aparato final. Funcionarían en grupos para a partir de las materias primas que se pudieran utilizar, dar origen a ramitas más largas. Y luego se unirían ellos mismos al resultado final. Los arbustos más grandes se unirían para dar lugar a ramas todavía mayores, y se repetiría el proceso hasta que un pequeño grupo se juntara para unirse a la raíz. De la misma manera, un arbusto que fuera suficientemente grande debería poder organizar los recursos que estén a su alcance para crear las diminutas semillas que dan origen al proceso (o, simplemente, repararse si se estropea, o crecer). Podría fabricar las partes más pequeñas utilizando técnicas parecidas a las de microacabado de los circuitos integrados actuales. Si el tamaño de las ramas más pequeñas fuera del orden de algunos átomos (es decir, la longitud de las mismas se podría medir en nanómetros), el robot arbusto podría coger átomos individuales de materia prima y unirlos, uno por uno, para formar partes nuevas. Esto sería una variación de los métodos de nanotecnología que se mencionaron en el Capítulo 2.



Un robot arbusto

Para concretar, podemos hacer un diseño real. Supongamos que la estructura básica es una rama grande que se divide en cuatro más pequeñas, cada una de ellas la mitad de la inicial. Si empezamos con una raíz que mida un metro y tenga diez centímetros de diámetro, y repetimos el proceso veinte veces, nos encontraremos con un arbusto de un billón de «hojitas», cada una de las cuales medirá una millonésima de metro (una micra) y tendrá un diámetro que será la décima parte de esa longitud. Como las hojas son mucho menores en peso y tamaño, se pueden mover a una velocidad un millón de veces

superior a la del tronco. Supongamos que el tronco puede oscilar hacia delante y hacia atrás una vez por segundo. Entonces, las hojas vibrarán un millón de veces por segundo. Si el arbusto se convierte en un haz apretado, la sección transversal será aproximadamente constante y tendrá una longitud de dos metros. El tronco ocupará la mitad de esta longitud, el segundo nivel de ramas la mitad de la longitud sobrante y así sucesivamente. Si se despliega y se deja con forma de paraguas, se convertirá en un disco de un diámetro algo menor a dos metros, tupido aunque poco denso cerca del centro y más ralo cerca del borde, con pequeños espacios vacíos en forma de huso, cuya punta es de una micra.

Si cada una de las uniones puede medir las fuerzas y los movimientos que se le aplican, entonces tendremos un excelente sensor. Contamos con un billón de «hojas dedo», cada una de las cuales puede detectar un movimiento de una décima de micra y una fuerza de algunos microgramos a velocidades de hasta un millón de oscilaciones por segundo. Es muy superior a la capacidad sensorial del ojo humano, que cuenta con un millón de puntos distinguibles, los cuales pueden detectar cambios que se produzcan, como mucho, cien veces por segundo. Si nuestro arbusto pone los dedos en una fotografía, puede «ver» la imagen con muchísimos detalles simplemente sintiendo las variaciones de peso de la plata del papel. Puede ver una película pasando los dedos por la cinta a gran velocidad. No existe ninguna razón para que no sea sensible a la luz, la temperatura y otros efectos electromagnéticos. De hecho, los dedos más pequeños tienen el tamaño adecuado para ser «antenas» para la luz. El arbusto podría tener un ojo si llevara una lente y colocara algunos millones de sus dedos en el plano focal situado tras ella. Incluso podría pasarse sin lente si colocara algunos dedos a cierta distancia, de manera que la luz se pudiera difractar formando, así, una lente holográfica.

Además de tener una capacidad sensorial semejante a la de los seres humanos que pueblan el mundo en la actualidad, nuestro arbusto podría modificar su entorno a la misma velocidad prodigiosa. Un ser humano bien entrenado, por medio de

movimientos corporales y manuales bien medidos, con los que puede cambiar de dirección, como mucho, varias veces por segundo con una precisión no superior a un pequeño porcentaje del movimiento total, podría modificar su entorno a una velocidad de miles de *bits* por segundo. Por ejemplo, una persona que escribe a máquina con mucha rapidez produce menos de cien *bits* de texto por segundo. La velocidad potencial de datos de un robot que tenga un billón de dedos, cada uno de ellos capaz de moverse un millón de veces por segundo, es más de mil billones (10^{15}) de veces mayor. Estas elevadas velocidades de datos suponen una tremenda coordinación de una enorme potencia de proceso, pero imaginemos las posibilidades. El robot arbusto podría ponerse en contacto con una pieza complicada de un delicado equipo mecánico —o, incluso, con un organismo vivo—, percibir simultáneamente la posición relativa de millones de partículas, algunas quizá del tamaño de las moléculas, y volverlas a ordenar para realizar una reparación casi instantánea. En la mayor parte de los casos, el superior sentido del tacto sustituiría totalmente a la visión, y la extrema habilidad eliminaría la necesidad de utilizar herramientas especiales.

Para controlar esta máquina maravillosa haría falta una cantidad astronómica de pensamiento. Gran parte de su funcionamiento estaría controlado por lo que en los animales se denomina *arcos reflejos*, pequeñas porciones del sistema nervioso situadas cerca del lugar que controlan. Cada una de las ramitas tendría la parte de un ordenador capaz de controlar la actividad de rutina. Solamente las situaciones excepcionales requerirían la intervención de los ordenadores de mayor tamaño situados cerca de la raíz. Si la rama tuviera también su propia fuente de energía (pensemos en una pila recargable) y posibilidades de comunicarse a distancia (radio o vibraciones sonoras de algunos cientos de cilios sincronizados, por ejemplo), el arbusto se podría convertir en multitud de pequeños arbustos coordinados entre sí. Cuanto menor fuera un arbusto, sería menos inteligente y tendría menos potencia. Pero su raíz madre lo podría

volver a programar y cargar para que realizara una cierta función y volviera, tan rápido como le fuera posible, a presentar su informe y recibir nuevas instrucciones.

Una de sus ventajas sería el tamaño: un robot pequeño se puede meter en espacios pequeños. En una máquina diminuta, la relación superficie/peso es mayor: un arbusto grande podría moverse por el suelo sin peligro, utilizando sus ramas como si fueran los ágiles dedos de los pies, pero una máquina más pequeña podría andar por el techo como una mosca, agarrándose con los cilios diminutos a las microscópicas grietas de la pintura o quedándose pegada a causa de la adhesión molecular. Los arbustos podrían cavar, aflojando partículas de tierra, y echándolas hacia atrás, y nadar, adoptando una forma aerodinámica y compacta, si los cilios formaran una piel capaz de bombear el fluido para impulsarse y fueran capaces de percibir el flujo, para evitar las turbulencias. Las máquinas extremadamente pequeñas tendrán tanta área superficial para su peso que podrán volar como los insectos, batiendo los cilios para mover el aire.

La contribución de las ramas mayores a la energía y a la inteligencia de las más pequeñas se puede visualizar como una especie de esquema piramidal invertido. Cada uno de los niveles de nuestro robot (contando todas las ramitas de ese nivel) tiene el doble de volumen que el siguiente nivel más pequeño y, en consecuencia, sitio para almacenar el doble de energía y un ordenador el doble de grande. Dos niveles más abajo, la relación es de cuatro a uno, y tres niveles más abajo, de ocho a uno. Si el control y la energía para cada nivel viene de las ramas que están situadas tres pasos más cerca de la raíz, las ramas pequeñas serán cuatro veces más vigorosas que de otra manera. Sólo saldrán perdiendo la raíz y las ramas que salen de ella. Como la mayor parte de su potencia y de su atención se dirigen hacia la parte superior del árbol, son incapaces de hacer muchos movimientos y sirven simplemente para proporcionar un marco estable si cierran pasivamente sus uniones.

Una cuestión importante es cómo funcionarán los programas de control de semejante bestia. En un caso extremo, podemos imaginar un programa que intente determinar la combinación de acciones de cada una de las uniones individuales que servirá para llevar a cabo la tarea deseada. Es, ciertamente, un ejemplo de un problema con un PND (polinomio no determinista) que se puede resolver solamente si se examinan todas las combinaciones de movimientos posibles y se elige la mejor (ver Apéndice 3). En la actualidad, estos problemas son irresolubles incluso para un manipulador que posea sólo un puñado (!) de dedos. Aunque los ordenadores serán mucho más perfectos en el futuro, el problema que plantea un sistema que cuenta con tantos dedos es infinitamente mayor. De vez en cuando, se descubrirá alguna estrategia especialmente ingeniosa para coordinar los miles de millones de dedos y hacer que desempeñen una tarea concreta. Y la gente reunirá estas estrategias en colecciones y se las intercambiará: serán las técnicas manuales de la era superinteligente. Pero mientras que los seres humanos enseñan cosas tan simples como la manera de atarse los cordones de los zapatos, las lecciones de las máquinas superinteligentes se podrán comparar a las instrucciones para montar un avión. Si bien será prácticamente imposible hallar la mejor solución para los problemas de rutina, no debería ser muy difícil encontrar una buena. Me imagino una estrategia de «divide y vencerás», en la que la raíz considera el problema globalmente y genera tareas adecuadas para los cuatro subárboles directamente conectados con ella. Éstos, a su vez, subdividen el problema y pasan los fragmentos a las ramas siguientes. El proceso se repite y las ramas más pequeñas reciben órdenes sencillas como moverse hasta una cierta posición o hasta que se encuentre resistencia. Una orden de la raíz podría ser la siguiente: *Arbusto Norte: quédate en la parte izquierda del plano A, a la derecha del plano B, y aplícale al objeto un vector fuerza total V; Arbusto Este: quédate a la derecha de Ay de B y oponte a cualquier movimiento que te aleje más de 10 cm del eje; Arbusto Sur: a la derecha de Ay ala izquierda de B, aplica una fuerza negativa V; Arbusto Oeste: a la izquierda de A y de*

B y oponte. Si un arbusto pequeño no puede resolver uno de los subproblemas tal y como le ha llegado, se quejará a la rama que se lo ha enviado y ésta intentará planificar algo mejor.

Contemplar un robot arbusto sería una maravilla de surrealismo. Sería diferente a cualquier cosa que se pudiera ver sobre la Tierra, a pesar de su parecido estructural con muchos seres vivos. Con su gran inteligencia, excelente coordinación, velocidad astronómica y enorme sensibilidad ante su entorno, podría hacer cosas sorprendentes constantemente y, al mismo tiempo, mantener su gracia perpetua. Los animales bípedos tienen tres o cuatro formas de andar efectivas. Los cuadrúpedos, algunas más. Los humanos con dos manos pueden coger un objeto de dos o tres formas. Un mecanismo con un billón de extremidades y con un gran cerebro es un ser que pertenece a otro orden. Añadamos a esto la capacidad de fragmentarse en una nube de diminutas moscas coordinadas y parecerá que las leyes de la física se someten a la intención y la voluntad. Alrededor de un robot arbusto, las cosas simplemente *sucederán*, algo que no ha podido hacer ningún tipo de magia. Imaginemos lo que debe ser habitar en un cuerpo semejante.

TRANSMIGRACIÓN

Algunos de nosotros tenemos concepciones del mundo bastante egocéntricas. Contamos con que, en el curso de nuestra existencia, se descubrirán métodos para alargar la vida humana y esperamos que, dentro de algunas eras, se pueda explorar el universo. Pero resulta decepcionante pensar que nuestra progenie artificial nos tratará con mucha altanería. La vida larga pierde bastante interés si estamos condenados a compartirla estúpidamente con nuestras máquinas superinteligentes, mientras ellas intentan describirnos sus descubrimientos, cada vez más espectaculares, mediante balbuceos infantiles para que los podamos entender. Lo que queremos es ser los actores principales de este juego superinteligente. ¿Qué posibilidades tenemos de serlo?

Una de las opciones puede ser la ingeniería genética. Las generaciones venideras de seres humanos se podrán diseñar utilizando las matemáticas, las simulaciones con ordenador y la experimentación, como se hace en la actualidad con los ordenadores, los aviones y los robots. Se podrán mejorar tanto sus cerebros como sus metabolismos y, de esta manera, podrán vivir cómodamente en el espacio. Pero es de suponer que seguirán formados de proteínas y de neuronas. Lejos de la Tierra, las proteínas no son un material ideal. Sólo son estables a unas ciertas temperaturas y presiones, son muy sensibles a la radiación y excluyen muchos componentes y técnicas de composición. Y no es muy probable que las neuronas, que no pueden cambiar más de mil veces por segundo, puedan llegar alguna vez a las velocidades de miles de millones por segundo de los componentes de los ordenadores de hoy en día. Antes de que pase mucho tiempo se fusionarán las tecnologías convencionales, miniaturizadas hasta la escala del átomo, y la biotecnología, con todas sus interacciones moleculares explicadas en términos mecánicos detallados, y formarán un conjunto sin fisuras de técnicas que abarcarán todos los materiales, tamaños y complejidades. Los robots estarán hechos de una mezcla de sustancias fabulosas entre las que se contarán, cuando sea necesario, materiales biológicos vivientes. En ese tiempo, un superhumano diseñado por medio de la ingeniería genética será algo así como un robot de segunda, ya que tendrá el inconveniente de que solamente se podrá construir un ser semejante a partir de la síntesis de proteínas derivadas del ADN. Para los fanáticos defensores del género humano, solamente tendrá una ventaja: muchas más limitaciones humanas que otros robots.

Los robots, tanto de primera como de segunda dejan nuestra pregunta sin contestar: ¿existe alguna posibilidad de que podamos compartir (usted y yo, personalmente) por completo este mundo mágico futuro? Esto exigiría un proceso que le confiriera al individuo todas las ventajas de las máquinas sin que perdiera su identidad. En la actualidad, mucha gente conserva la vida porque existe un arsenal de órganos y de otras partes del cuerpo hechos artificialmente. Con el tiempo, sobre todo a

medida que vayan perfeccionándose las técnicas robóticas, estas piezas de repuesto serán mejores que las originales. Entonces, ¿por qué no sustituirlo todo, es decir, trasplantar un cerebro humano a un cuerpo de robot diseñado especialmente? Por desgracia, aunque esta solución pudiera vencer nuestras limitaciones físicas, no afectaría para nada a nuestro inconveniente principal: la inteligencia limitada del cerebro humano. El trasplante sacaría el cerebro de nuestro cuerpo. ¿Existe alguna forma de sacar nuestra mente de nuestro cerebro?

Le acaban de meter en un quirófano. Está de servicio un robot cirujano del cerebro. Usted tiene a su lado un ordenador que espera convertirse en un equivalente humano. Lo único que le falta para empezar a funcionar es un programa. Le anestesian el cráneo, pero no el cerebro. Usted está totalmente consciente. El robot cirujano le abre el cráneo y coloca una mano sobre la superficie del cerebro. Esta mano poco corriente contiene maquinaria electrónica y está conectada, por medio de un cable, al ordenador móvil que tiene usted a su lado. Los instrumentos de la mano examinan los primeros milímetros de la superficie del cerebro. Se llevan a cabo medidas de resonancia magnética de alta resolución y se elabora un mapa químico tridimensional al mismo tiempo que una multitud de antenas eléctricas y magnéticas recogen señales y las interpretan rápidamente para revelar, momento a momento, los impulsos que se trasmiten entre las neuronas. Estas medidas, junto con el conocimiento global de la arquitectura neurológica del ser humano, le permiten al cirujano escribir un programa que modele el comportamiento de la capa superior del tejido cerebral examinado. Se instala el programa en una pequeña porción del ordenador que está esperando y se activa. Las medidas hechas con la mano le proporcionan copias de las entradas que recibe el tejido original. El cirujano y usted comprueban la exactitud de la simulación comparando las señales que produce con las originales. Se trasmiten con mucha rapi-

dez, pero todas las diferencias aparecen en pantalla. El cirujano ajusta la simulación hasta que la correspondencia sea casi perfecta.

Para que a usted no le quepa ninguna duda sobre la exactitud de la simulación, tiene un botón que puede pulsar si desea comparar por sí mismo la simulación con el funcionamiento del tejido original. Cuando aprieta el botón, se activa una multitud de electrodos situados en la mano del cirujano. Estos electrodos, por medio de inyecciones muy precisas de corriente y de impulsos electromagnéticos, pueden anular la actividad normal de emisión de señales de las neuronas próximas. Están programados para poder inyectar la salida de la simulación en aquellos lugares en los que el tejido simulado señala a otros sitios. Cuando usted aprieta el botón, una simulación del ordenador sustituye a una pequeña parte de su sistema nervioso. Usted aprieta el botón, lo suelta y lo vuelve a apretar. No siente ninguna diferencia. En cuanto usted queda satisfecho, se establece permanentemente la conexión de la simulación. El tejido del cerebro está ahora indefenso: recibe las entradas y reacciona de la misma forma que antes, pero se ignora la salida. Los manipuladores microscópicos situados en la superficie de la mano separan las células del tejido superfluo y las pasan a un aspirador, donde se las elimina.

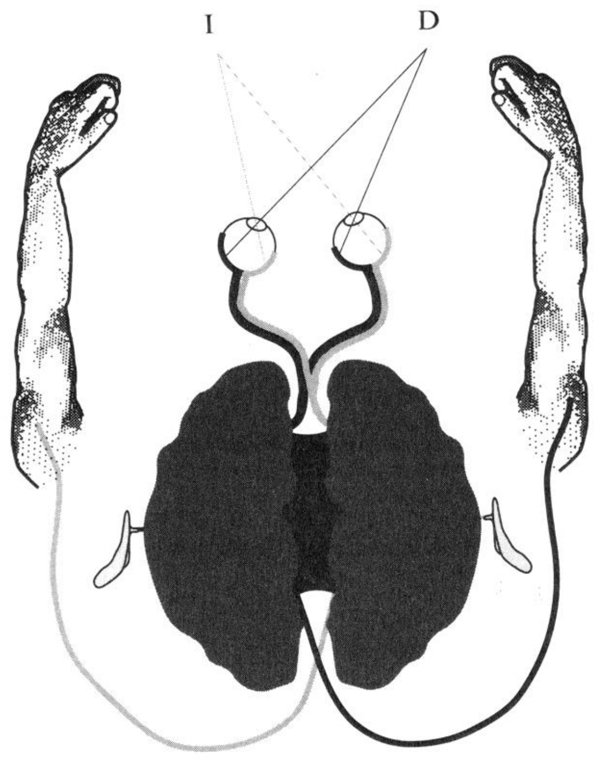
La mano del cirujano penetra una fracción de milímetro más en su cerebro y compensa instantáneamente todas las medidas y señales para la nueva posición. Se repite el proceso en la capa siguiente y en seguida aparece una segunda simulación en el ordenador, que se comunica con la primera y con el resto del tejido cerebral original. A medida que avanza el proceso, se va simulando y excavando en el cerebro, capa tras capa. Finalmente, su cráneo se queda vacío y la mano del cirujano se apoya en la base del mismo. Aunque usted no ha perdido la conciencia, ni siquiera el hilo de sus pensamientos, su mente ha pasado de su cerebro a una máquina. El paso final y un tanto desconcertante es cuando el cirujano retira la mano. Su cuerpo, repentinamente abandonado, sufre algunos espasmos y muere. Durante un momento, lo único que experimenta

usted es quietud y oscuridad. Y, de nuevo, puede abrir los ojos. Pero ahora su perspectiva ha cambiado. La simulación del ordenador se ha desconectado del cable que la unía con la mano del cirujano y se ha conectado con un cuerpo nuevo y brillante, del estilo y color que usted había elegido. Su metamorfosis ha terminado.

Existen otras formas de llevar a cabo la transferencia de la mente humana a una máquina, para los que sean aprensivos. Un examen cerebral de alta resolución puede fabricar un nuevo «Usted» de un solo golpe y sin cirugía. Si incluso esta última técnica le resulta demasiado «invasora», podemos imaginar un método más psicológico. Consiste en programar una especie de ordenador portátil (acaso como las gafas mágicas) con las constantes universales de la mentalidad humana, su estructura genética y todos los detalles posibles sobre su vida. Este ordenador llevará un programa que le permitirá hacer excelentes imitaciones. Usted lo llevará consigo durante los mejores años de su vida, y él escuchará y observará todo. Quizá controle su cerebro y aprenda a anticipar todos los movimientos y respuestas que usted pueda hacer o dar. Pronto podrá imitarle convincentemente por teléfono y engañar a sus amigos. Cuando usted muera, el programa se instalará en un cuerpo mecánico que asumirá su vida y sus responsabilidades.

Si usted es vertebrado, existe otra opción que combina las características de venta de todos los métodos que acabamos de describir. El cerebro de los vertebrados consta de dos hemisferios conectados entre sí por varios haces grandes de fibras nerviosas. La mayor se denomina *corpus callosum*. En los años sesenta, los investigadores de California, después de llevar a cabo experimentos con animales, curaron a algunos enfermos que padecían tipos de epilepsia resistentes a los tratamientos cortándoles los *corpora callosa*. (¡Los robots médicos del futuro no utilizarán términos latinos!) Aunque parezca mentira, en un principio este método no producía efectos secundarios en los pacientes. El *corpus callosum*, con sus doscientos millones de fibras nerviosas, es la interconexión a larga distancia más

grande del cerebro. Es mucho más grueso que el nervio óptico y que la médula espinal. Si se corta el nervio óptico, la víctima se queda completamente ciega. Si se secciona la médula espinal, el cuerpo se queda fláccido e insensible. Pero si se corta el enorme cable que une los dos hemisferios no pasa nada. Bueno, casi nada. Si se trasmite el nombre de un objeto, «cepillo», por ejemplo, a la parte derecha del campo visual de una persona con el «cerebro escindido», ésta no lo puede seleccionar de entre un grupo de varios objetos con la mano izquierda, pero no tiene ningún problema si lo hace con la derecha. A veces, en la versión de la mano izquierda del ejercicio, la mano derecha se mueve aparentemente enfadada para guiar a la izquierda al sitio exacto.



◄ El corpus callosum

La mitad izquierda del cerebro controla la mitad derecha del cuerpo y viceversa. La mitad izquierda, por lo general, se dedica al lenguaje y a los cálculos, mientras que la mitad derecha está especializada en el razonamiento espacial. Las dos mitades del cerebro se comunican normalmente por medio del *corpus callosum*, pero pueden continuar funcionando independientemente si éste se secciona.

La neuroanatomía sugiere alguna explicación a este fenómeno. Los nervios que controlan los músculos de la izquierda del cuerpo y las zonas del nervio óptico que contemplan la izquierda de la escena visual están conectados solamente con la parte derecha del cerebro. De la misma manera, la zona izquierda del cerebro controla la derecha del cuerpo y contempla la mitad derecha del campo visual, como se puede ver en la figura de la página 134. Normalmente, las dos mitades funcionan en íntima colaboración y la información que descubre una de ellas pasa rápidamente a la otra por medio del *corpus callosum*. En una persona con el cerebro escindido, no se produce este flujo de información. Las dos mitades tienen que descubrir las cosas independientemente. La mitad de la derecha no sabe lo que hace la de la izquierda. Sin embargo, parece que las dos son conscientes de las emociones de la otra y que esta información proviene aparentemente de las conexiones nerviosas de la base del cerebro, que han quedado intactas.

Roger Sperry, del Instituto de Tecnología de California, que recibió el Premio Nobel de 1981 por sus descubrimientos sobre la función del *corpus callosum*, descubrió que en los sujetos con el cerebro escindido cada una de las dos mitades parece tener una personalidad humana inteligente, totalmente consciente e independiente de la otra. Se sabe que en los cerebros intactos algunas de las fibras del *corpus callosum* realizan funciones básicas, tales como recombinar las dos mitades de los campos visuales de los ojos, mientras que otras deben comunicar entre los dos hemisferios conceptos mentales más elevados. Es decir, tenemos motivos para creer que el *corpus callosum* es como una ventana amplia y bien organizada para contemplar las actividades mentales de los dos hemisferios.

Supongamos que en el futuro, cuando se conozca mejor la función del cerebro, le seccionan su *corpus callosum* y se conectan los extremos cortados con un ordenador externo. El ordenador está programado, en un principio, para pasar las informaciones de un hemisferio al otro y para «fisgonearlas». A partir de las cosas que se entera, construye un modelo de sus actividades mentales. Después de un cierto tiempo, empieza a introducir en ese flujo sus propios mensajes, con lo que se va insinuando gradualmente en su pensamiento y le va proporcionando nuevos conocimientos y capacidades. Con el tiempo, cuando su cerebro se vaya consumiendo por la edad, el ordenador irá asumiendo las facultades que usted va perdiendo. Finalmente, su cerebro morirá y su mente habrá pasado entera al ordenador. Puede que, debido a los avances del *scanning* de alta resolución, se pueda conseguir esto sin necesidad de recurrir a la cirugía. Lo único que tendrá usted que hacer es ponerse algún casco o banda en la cabeza para que controle y modifique el tráfico entre los hemisferios por medio de campos electromagnéticos cuidadosamente controlados.

MUCHOS CAMBIOS

Sea cual sea el estilo de transferencia mental que usted elija, muchas de sus antiguas limitaciones desaparecerán a medida que se vaya produciendo el proceso. Su ordenador tiene un mecanismo de control llamado «velocidad». Se ha fijado en «lento», para que las simulaciones estén sincronizadas con el cerebro antiguo, pero ahora usted lo pasa a «rápido», lo que le permite comunicarse, reaccionar y pensar mil veces más de prisa. Se puede copiar todo este programa en distintas máquinas, con lo que tenemos dos o más versiones pensantes de usted. A lo mejor usted desea cambiar su mente de un ordenador a otro que sea más avanzado o se adapte mejor al nuevo entorno. Asimismo, se puede copiar el programa en un equivalente de lo que será una cinta magnética del futuro. Si la má-

quina en la que usted vive se estropea sin posibilidad de arreglo, se puede pasar la cinta a un ordenador vacío, y el resultado será usted menos sus experiencias desde que se hizo la copia. Si se hacen muchas copias y no se dejan juntas, es bastante improbable que se produzca su muerte permanente.

Su mente, lo mismo que un programa de ordenador, puede viajar por los canales de información codificada, por ejemplo, como mensaje en un rayo láser que se emite entre los planetas. Si usted descubre que hay vida en una estrella de neutrones y quiere hacer un viaje hasta ella, debe inventar una manera de construir allí un robot de material neutrónico y transferirle su mente. Como las reacciones nucleares son aproximadamente un millón de veces más rápidas que las químicas, el usted neutrónico debe ser capaz de pensar con una velocidad un millón de veces superior. Podrá explorar, tener experiencias y adquirir nuevos recuerdos y, luego, emitirse de nuevo a casa. Su cuerpo original puede mantenerse aletargado durante el viaje, y se reactivará, junto con los nuevos recuerdos, cuando llegue el mensaje de regreso. Puede que tarde un minuto, pero llevará las experiencias de todo un año subjetivo. También lo puede mantener en actividad. Y habría dos versiones distintas de usted, con recuerdos diferentes, a causa del intervalo del viaje.

Las cosas que ha aprendido pueden implicar cambios en su personalidad. Muchos de los cambios nacerán de los juegos que usted realice con su antiguo programa. Cuando haya acelerado mil veces su velocidad de control se dará cuenta de que dispone de varias horas (subjetivamente hablando) para reaccionar ante situaciones que antes exigían respuestas instantáneas. Mientras un objeto se está cayendo, usted tiene tiempo para sopesar las ventajas e inconvenientes de intentar cogerlo al vuelo, incluso para resolver las ecuaciones diferenciales de su movimiento. Si se encuentra en una situación social difícil, tendrá tiempo para leer un libro sobre buenos modales y decidir qué es lo que puede hacer. Si tiene que reparar una máquina averiada, antes de ponerle la mano encima podrá aprenderse su teoría de funcionamiento y pensar cuidadosamente cuáles son las cosas que fallan. Hablando en general, tendrá tiempo para

aplicar los hoy considerados métodos de investigación avanzada para resolver triviales problemas cotidianos.

Tendrá tiempo pero, ¿tendrá también paciencia, o ese aumento de su velocidad mental simplemente le paralizará de aburrimiento? El aburrimiento es un mecanismo mental que le evita perder el tiempo apartándole de actividades inútiles, pero si actúa demasiado pronto o con demasiada agresividad limita su capacidad de atención y, por lo tanto, su inteligencia. Uno de los cambios que deseará hacer en su programa es retrasar la aparición del aburrimiento hasta más allá del límite que se ha descubierto que tienen las inteligencias más preclaras. Cuando lo haya hecho, se dará cuenta de que trabaja muy a gusto en la resolución de largos problemas con muchísimas ramificaciones. De hecho, sus pensamientos se implican tanto y de forma tan rutinaria, que tiene que desarrollar su memoria a corto plazo. También tendrá que hacerlo con su memoria a largo plazo, ya que un mes de acontecimientos puede ocupar el tiempo objetivo de un siglo. Y éstos no son más que los primeros cambios.

Antes he mencionado la posibilidad de hacer copias de uno mismo y que cada una de ellas corra sus propias aventuras. También sería posible fundir los recuerdos de las distintas copias. Para evitar la confusión, los recuerdos de los acontecimientos indicarían a qué cuerpo pertenecen, lo mismo que nuestros recuerdos suelen estar en un contexto que establece la época y el lugar del suceso recordado. Y se podría hacer esta fusión no sólo con recuerdos de dos versiones de la misma persona, sino con distintas personas. Las fusiones selectivas, es decir, con ciertos recuerdos de otra persona y no con otros, serían una forma superior de comunicación y se podrían compartir los recuerdos, las habilidades, las actitudes y las personalidades de manera rápida y efectiva. Su cuerpo nuevo tendrá muchos más recuerdos que el biológico original, pero la explosión de información acelerada hará que sea imposible atesorar todos los conocimientos de la civilización. Usted tendrá que elegir y seleccionar el contenido de su mente. Con frecuencia tendrá acceso a conocimientos y técnicas hasta ahora reservadas a

mentes superiores a la suya, y será abrumadora la tentación de sustituir esas aptitudes por las suyas. A largo plazo, podrá recordar muchas experiencias de otras personas y se encontrarán en mentes ajenas los recuerdos que usted originó. Algunos conceptos, tales como vida, muerte e identidad, perderán su significado actual a medida que se combinan, mezclan y recombinan sus fragmentos mentales y los de otras personas y forman asociaciones temporales, unas veces grandes y otras pequeñas, muchas veces aisladas, otras altamente individuales y otras efímeras, simples olas de los rápidos del torrente de conocimientos de la civilización. Ya existen muestras de este tipo de fluidez. Los seres humanos, por medio de la cultura, aprenden técnicas y aptitudes de otros durante toda su vida. Genéticamente, en los grupos sexuados, cada individuo independiente es un conjunto temporal de genes que se combinan y recombinan de forma diferente en cada generación.

Las transferencias mentales no tienen por qué limitarse a los seres humanos. Existen otras especies sobre la Tierra que tienen cerebros grandes, desde los delfines, cuyo sistema nervioso es tan grande y complejo como el nuestro, hasta los elefantes, ballenas y calamares gigantes, cuyos cerebros pueden llegar a tener un tamaño veinte veces superior al nuestro. Aunque no se sabe exactamente qué tipos de mentes y culturas tienen estos animales, el tema sigue siendo objeto de controversia. Pero la historia de su evolución es tan larga como la nuestra y, con toda seguridad, en sus estructuras cerebrales y en sus recuerdos tienen codificadas genéticamente muchas informaciones únicas y que les ha costado mucho esfuerzo conseguir. Los métodos de transferencia cerebro-a-ordenador que funcionen para los seres humanos funcionarán también para estos animales de cerebros grandes, y eso nos permitirá entretejer en nuestro tapiz cultural sus pensamientos, capacidades y motivaciones. Con otros métodos ligeramente diferentes, más centrados en la genética y en la estructura física que en la vida mental, podremos almacenar en bancos de datos todas las informaciones contenidas en otros seres vivos de sistema nervioso pe-

queño o que carezcan de él. Los organismos más sencillos pueden contribuir con poco más que la información contenida en su ADN. De esta manera, nuestros congéneres futuros podrán aprovecharse de lo que ha aprendido la biosfera terrestre a lo largo de su historia de miles de millones de años. Y estos conocimientos estarán más seguros si se guardan en bancos de datos repartidos por todo el universo. De acuerdo con los esquemas actuales, los genes y las ideas se pierden con frecuencia en nuestra frágil y pequeña Tierra, cada vez que cambian las condiciones que los originaron.

Nuestras especulaciones terminan con una supercivilización, la síntesis de toda la vida del sistema solar, que se perfecciona y extiende constantemente, llega más allá del sol y transforma la no vida en mente. Es muy posible que, en otros lados, existan otras burbujas en expansión. ¿Qué sucederá si nos encontramos con una de ellas? Una de las posibilidades es una fusión negociada que solamente requerirá un esquema de equivalencia entre las representaciones de la memoria. Este proceso, que posiblemente esté sucediendo en alguna parte, convertiría a todo el universo en un gran ente pensante, preludio de cosas todavía más maravillosas.

¿QUÉ ES LO QUE SOY YO?

La idea de que la mente humana se pueda transferir a un cuerpo nuevo a veces se tiene que enfrentar con una fuerte objeción que presentan personas que no discuten su posibilidad teórica: «Sea cual sea el procedimiento para hacer la copia, el resultado final será una persona distinta. Si soy yo el ser que se está copiando, la copia, aunque se crea que soy yo, es simplemente un impostor que se engaña a sí mismo. Si el proceso de copiado destruye el original, entonces es que me han matado. Y la copia se puede divertir muchísimo explorando el universo con mi nombre y mis habilidades, pero eso no le proporcionará ningún consuelo a mis restos mortales.»

Este punto de vista, al que llamo *postura de identidad con el cuerpo,* hace que la ampliación de la vida por duplicación sea mucho menos interesante, personalmente hablando. Creo que se puede salvar la objeción si se acepta, como alternativa, otra postura a la que llamo *de identidad con el modelo*. La identidad con el cuerpo supone que una persona se define por el material del que está hecho el cuerpo humano. Solamente podemos conservar a la persona como individuo si mantenemos la continuidad de ese material corporal. Por el contrario, la identidad con el modelo define la esencia de una persona, yo por ejemplo, como el *modelo* y el *proceso* que tienen lugar en mi cabeza y en mi cuerpo, no como la maquinaria en la que se lleva a cabo ese proceso. Si se puede conservar el proceso, yo estaré a salvo. El resto es paja.

Creo que la postura de identidad con el cuerpo se basa en una intuición errónea sobre la naturaleza de las cosas vivas. De una manera sutil, la conservación del modelo y la pérdida de la sustancia es normal en la vida cotidiana. Lo mismo que nosotros, seres humanos, comemos y excretamos, las células viejas de nuestros cuerpos mueren, se expulsan y se sustituyen por copias hechas de materiales nuevos. Los pocos componentes del cuerpo que tienden a ser más estáticos, como las células nerviosas, tienen metabolismos que hacen que, poco a poco, se vayan sustituyendo sus partes interiores. Hacia la mitad de la vida es necesario reemplazar todos los átomos que están en nuestro cuerpo cuando nacemos. Solamente nuestro modelo, y solamente parte de él, permanece con nosotros hasta nuestra muerte.

Quisiera explorar algunas de las consecuencias de la postura de la identidad con el modelo. Los transmisores de materia han aparecido con frecuencia en la literatura de ciencia ficción, al menos desde que se inventaron las máquinas para hacer reproducciones exactas o facsímiles, a finales del siglo XIX. Sugiero la idea solamente como experimento mental, para simplificar algunos de los puntos de mi propuesta sobre la transferencia de la mente. Un transmisor facsímil examina una fotografía línea por línea con una fotocélula sensible a la luz y produce

una corriente eléctrica que varía según el brillo del punto examinado de la foto. Esta corriente eléctrica variable se transmite por medio de alambres hasta un punto lejano, donde controla la luminosidad de una bombilla en un receptor facsímil. El receptor examina la bombilla sobre el papel fotosensible de la misma manera que el transmisor. Cuando se revela el papel, se obtiene un duplicado de la fotografía original. Este dispositivo representó un gran adelanto para los periódicos, ya que podían recibir casi instantáneamente ilustraciones de partes lejanas del país, en vez de tener que esperar a que las placas fotográficas llegaran por tren.

Si se puede hacer con fotografías, ¿por qué no con objetos sólidos? Un *transmisor de materia* puede examinar un objeto e identificar sus átomos o moléculas de una en una y, posiblemente, llevárselas en el curso del proceso. La identidad de los átomos se transmitiría a un receptor donde se podría montar un duplicado del original, en el mismo orden, siempre y cuando hubiera allí existencias de átomos. Los problemas técnicos son muy grandes, pero es fácil entender el principio básico, como atestiguarán todos los aficionados a *Star Trek*.

Si se puede hacer con objetos sólidos, ¿por qué no con una persona? Lo único que hay que hacer es «conectarla» al transmisor, encender el examinador *(scanner)* y saludarla cuando salga del receptor. Pero ¿será realmente la misma persona? Si el sistema funciona bien, el duplicado será totalmente indistinguible del original. Sin embargo, supongamos que se le olvida conectar el receptor durante el proceso de transmisión. El transmisor examinará y desmontará a la víctima y enviará un mensaje que no se recibe a un receptor desconectado. La persona original morirá. ¿No matará el proceso a la persona original, esté conectado o no el receptor? ¿No es el duplicado sólo eso, una simple copia? O bien supongamos que haya dos receptores que captan el mensaje. ¿Cuál de los dos duplicados es el original auténtico, en caso de que lo fuere alguno?

La postura de la identidad con el cuerpo es muy clara por lo que se refiere a este punto. Un transmisor de materia es un complicado mecanismo de ejecución que le matará a usted y

colocará a un astuto impostor en su lugar. La postura de identidad con el modelo ofrece una perspectiva distinta. Supongamos que yo entro en la cámara de transmisión. El transmisor me examina y desmonta mi cuerpo gelatinoso, pero mi modelo (¡yo!) se mueve continuamente: sale de la gelatina que se está disolviendo, pasa por el haz transmitido y termina en la otra masa de gelatina que está en el punto de destino. En ningún momento se ha destruido al modelo (yo). Pero, ¿qué pasa con la cuestión de los duplicados? Supongamos que el transmisor de materia está conectado a dos receptores en vez de a uno solo. Después de la transferencia, habrá una copia de mí en cada uno de ellos. Por supuesto, por lo menos una de ellas es simplemente una copia. Yo no puedo ser las dos, ¿no es cierto? *¡Error!*

La suposición de que a cada persona le corresponde únicamente un cuerpo tiene sus raíces en toda nuestra experiencia anterior. Pero cuando pensamos en la posibilidad de separar la mente de la materia, de almacenarla y transmitirla, esta identificación simple, natural y evidente se convierte en algo confuso y engañoso. Consideremos el mensaje: «Yo no soy de gelatina.» Cuando lo escribo, sale de mi mente, va al teclado de mi ordenador y atraviesa miríadas de circuitos electrónicos y grandes cantidades de alambres. Después de aventuras sin cuento, el mensaje aparece en montones de libros como el que usted tiene entre manos. ¿Cuántos mensajes había? Afirmo que lo más útil es pensar que sólo había uno, a pesar de que se haya reproducido masivamente. Lo vuelvo a repetir: «Yo no soy de gelatina», y sigue habiendo solamente un mensaje. No habrá un segundo mensaje hasta que no introduzca una modificación significativa, por ejemplo «No soy de mantequilla». Y el mensaje no se destruirá hasta que no se pierda la última versión y hasta que no se difumine en la memoria de la gente y sea imposible reconstruirlo. El mensaje es la información que comunica, y no el medio en el que se ha codificado. El «modelo» que afirmo que es el yo real tiene las mismas propiedades que este mensaje. Y si se hace una copia momentánea de mi estado, sea en una cinta o en otro cuerpo, no se obtienen dos personas.

El «proceso» del aspecto es algo más complicado. En el instante en que se monta una «persona mensaje», no es más que una copia del original. Pero si existen dos copias, con el tiempo se convertirán en dos personas diferentes. Y el problema de hasta dónde tiene que llegar esta diferenciación para que la sociedad les confiera dos identidades únicas es tan espinoso como las preguntas «¿cuándo se convierte un feto en persona?», o «¿cuándo se puede considerar una especie diferenciada de una especie que está en proceso de evolución?». Pero si esperamos un tiempo cero, entonces las dos copias son la misma persona, y si destruimos una de ellas la persona seguirá existiendo en la otra copia. Todas las cosas que esa persona podría haber hecho o pensado siguen siendo posibles. Si, por el contrario, dejamos que las dos copias lleven vidas independientes durante un año y entonces destruimos una de ellas, entonces hemos matado a un ser humano único. Pero si esperamos un ratito antes de destruir una de las copias, entonces sólo se habrá perdido una pequeña cantidad de información única.

Este razonamiento podría ser un gran consuelo para usted en el caso de que estuviera a punto de enfrentarse con un gran peligro y supiera que se acababa de hacer una copia suya en cinta. Si usted muriese, se podría hacer una copia activa de la cinta. Sería ligeramente distinta de la versión de usted que había muerto, ya que faltarían los recuerdos desde el momento en que se copió. Pero un poco de amnesia resulta bastante trivial si se compara con la pérdida total de la memoria y del resto de las funciones, que es el resultado de la muerte si no existe copia.

El hecho de aceptar intelectualmente que existe un sustituto suyo reciente no le impedirá sentir inmensos deseos de conservación si se tiene que enfrentar con una muerte inminente, aunque sea por una buena causa. Estos sentimientos serían como un resto evolutivo de su pasado con una sola copia, de la misma manera que el miedo a volar es la respuesta a los porcentajes de accidentes aéreos. Los instintos antiguos no se borran automáticamente cuando cambian de repente las normas de la vida.

La postura de la identidad con el modelo tiene implicaciones dualísticas evidentes, ya que permite que la mente se separe del cuerpo. Aunque la mente no es otra cosa que la consecuencia de la interacción de la materia, la posibilidad de copiarla de un medio en el que está almacenada en otro le confiere una independencia y una identidad que la hace distinta de los mecanismos que hacen funcionar el programa. Este dualismo queda mucho más claro si pensamos en las diversas posibilidades de codificación.

Algunos superordenadores tienen miríadas de ordenadores independientes conectados entre sí por medio de una red que permite el libre flujo de informaciones entre ellos. El sistema operativo que se le aplique debe permitir que los procesos individuales pasen de un procesador a otro cuando están teniendo lugar, es decir, debe ser como una especie de juego de manos que permite que haya más procesos que procesadores. Si se instalara una mente humana en una máquina del futuro de este tipo, las funciones que en un principio realizaban ciertas partes del cerebro quedarían codificadas en un proceso específico. El juego de manos aseguraría que las operaciones que tuvieran lugar en zonas fijas del cerebro original pasaran rápidamente de un punto a otro de la máquina. Si el ordenador estuviera funcionando con otro programa, la simulación pasaría a grupos de procesadores totalmente diferentes. El proceso del pensamiento no se interrumpiría aunque cambiaran continuamente su localización y su maquinaria física, ya que se mantendría el modelo.

La forma más directa de llevar a cabo operaciones complejas, como las que tienen lugar en la mente humana, no suele ser la más efectiva. Un proceso que se puede describir como una larga secuencia de pasos de un programa a veces se puede transformar matemáticamente en otro que llega a la misma conclusión con muchas menos operaciones. Cuando era joven, el famoso matemático Carl Friedrich Gauss se distinguió en la escuela por ser muy rápido haciendo cálculos. Una vez, por diversión, uno de los profesores le propuso el problema de sumar todos los números entre 1 y 100. Dio la respuesta correcta en

menos de un minuto. Gauss había observado que los cien números se podían agrupar en cincuenta parejas, 1 + 100, 2 + 99, 3 + 98, 4 + 97, etc., y que cada una de las parejas sumaba 101. Así que multiplicó 50 por 101 y obtuvo la solución, 5.050, sin tener que hacer un montón de sumas aburridísimas. En los procesos de los ordenadores se pueden alcanzar estas velocidades. Los denominados *compiladores optimizados* tienen varios programas para acelerar las transformaciones. Algunos las aceleran muchísimo, y, en consecuencia, también los programas que traducen. La clave puede ser una reorganización total del orden de los cálculos y de la representación de los datos.

Existe un tipo de transformaciones, muy rápidas, para las que se necesita una serie de valores. Estos valores, por medio de la transformación, se combinan y dan lugar a otra serie tal que cada elemento de la misma es un número formado por una combinación única de todos los elementos de la serie primitiva. De esta manera, una sola operación con un elemento de la serie nueva sustituye, con frecuencia, a una multitud de operaciones con la serie original, y la eficacia es mucho mayor. También existen transformaciones análogas para la variable tiempo: se puede cambiar toda una secuencia de operaciones por otra equivalente en la que cada paso nuevo requiere una pequeña fracción del trabajo que se precisaba en la serie original. Lo concreto se vuelve difuso, y lo difuso concreto. Unos cuantos matemáticos pueden reescribir un programa y alterarlo hasta que quede irreconocible. Con un multiprocesador, los sucesos simples de la formulación original pueden aparecer, en el programa transformado, como correlaciones entre sucesos acaecidos en máquinas remotas y en tiempos remotos. Sin embargo, en sentido matemático, los cálculos transformados son exactamente iguales que los originales.

Así, si tuviéramos que transformar un programa que simulara una persona, la persona permanecería intacta: su mente es la propiedad matemática abstracta común a los dos programas, el transformado y el original, y no depende de la forma del programa. La mente, tal y como la he definido, no sólo no está ligada a un cuerpo en particular, sino que tampoco lo está a un

modelo en particular. Se la puede representar por medio de una clase infinita de modelos que sean equivalentes entre sí de una cierta forma, muy abstracta. Esta observación me tienta a hacer la breve extrapolación filosófica que se encuentra en el Apéndice 3.

La inmortalidad, tal y como la he descrito, es sólo una defensa temporal contra la pérdida inútil del conocimiento y de la función, el peor aspecto de la muerte personal. A largo plazo, para nuestra supervivencia será necesario que se produzcan cambios que no podemos elegir nosotros mismos. Tendremos que renunciar a algunas de nuestras partes y sustituirlas por otras para seguir el paso de las condiciones cambiantes y de la evolución de nuestros competidores. La supervivencia supone participar en una especie de Juegos Olímpicos cósmicos: en cada edición habrá nuevos acontecimientos y se batirán récords que acabarán con las viejas marcas. Con el tiempo, cada uno de nosotros será un ser completamente diferente y nuestra forma la dictarán las condiciones externas más que nuestros propios deseos. Nuestros recuerdos e intereses actuales perderán toda su importancia y terminarán en un polvoriento archivo que acaso consulten los historiadores en alguna ocasión. La muerte personal, tal y como la conocemos, diferirá de esta inevitabilidad sólo en su brusquedad relativa. Si se considera así, la inmortalidad personal por medio del trasplante de la mente es una técnica cuyo beneficio principal es mimar la sensibilidad y el sentimentalismo de los seres humanos. Me parece que nuestra civilización evolucionará en este mismo sentido, aunque no lleguemos a trasplantar la mente a los robots.

El individuo ancestral está siempre condenado a muerte y va dilapidando su herencia cuando trata de adaptarse a corto plazo a las posibilidades ambientales. Pero este proceso evolutivo, desde un punto de vista más positivo, significa que ya somos inmortales y que lo hemos sido desde los albores de la vida. Nuestros genes y nuestra cultura pasan continuamente de una generación a la siguiente, sometidos solamente a ligeras alteraciones para adaptarse a las continuas exigencias de los nuevos récords de los Juegos cósmicos. E incluso por lo que se

refiere a nuestra vida personal, ¿quién de nosotros desearía permanecer sin experimentar cambios, tener durante toda la vida los mismos conocimientos, recuerdos, pensamientos y habilidades que teníamos de pequeños? Los seres humanos valoran mucho los cambios y el desarrollo, y nuestros descendientes artificiales compartirán con nosotros este aprecio: tanto su supervivencia como la nuestra depende de ellos.

DESPERTAR EL PASADO

La posibilidad de trasplantar mentes facilitará la tarea de devolver a la vida a alguien que haya sido grabado y almacenado cuidadosamente. Pero, ¿qué sucederá si se pierde parte de la transcripción? A partir de otras informaciones se podrían reconstruir muchas de las partes perdidas: por ejemplo, el código genético de la persona, películas de su vida, muestras de escritura, historiales médicos o recuerdos de sus compañeros. Todo eso sería posible en un mundo de superinteligencias que tuvieran astronómicos poderes de observación y deducción. La postura de identidad con el modelo supone que una persona reconstruida por deducción sería tan real como otra reconstruida a partir de una cinta intacta. La única diferencia es que, en el primer caso, una parte del modelo de la persona habría quedado parcialmente difusa en el medio ambiente antes de que se llevara a cabo la reconstrucción.

Pero ¿qué sucedería si no existiera ninguna cinta? Los arqueólogos de la actualidad hacen deducciones verosímiles sobre figuras históricas a partir de fragmentos de documentos antiguos, pedazos de cerámica, exámenes de cuerpos momificados por medio de rayos X, hechos históricos demostrados, el conocimiento general de la naturaleza humana y todo lo que pueden encontrar. Los autores de ficción histórica utilizan estos datos para construir guiones detallados de cómo debieron suceder ciertos hechos. Los arqueólogos superinteligentes, pertrechados con instrumentos maravillosos (capaces, por ejemplo, de hacer medidas a escala atómica de objetos enterrados

profundamente), podrían desarrollar este proceso y ser capaces de reconstruir con todo detalle y en cualquier fase de su vida a personas muertas mucho tiempo antes.

La resurrección total se podrá poner en práctica por medio de inmensos simuladores. En la actualidad se utilizan ordenadores muy potentes para predecir el trayecto de los planetas y de las naves espaciales. La trayectoria precisa que siguió el Voyager 2 al pasar por Júpiter, Saturno, Urano y sus lunas y Neptuno se calculó por medio de simulaciones repetidas en las que se ensayaron distintos tiempos iniciales, direcciones y velocidades de la nave hasta que se encontró la combinación correcta. Con menos exactitud, los programas meteorológicos modernos simulan la acción de la atmósfera sobre el globo terrestre. Y se están empezando a verificar por medio de simuladores los nuevos diseños de aviones, las explosiones nucleares y otras muchas cosas. Estas simulaciones, que nos permiten mirar a hurtadillas los posibles futuros, nos dan la posibilidad de elegir entre ellos. Como las leyes de la física son simétricas en el tiempo, los simuladores se pueden accionar tanto hacia adelante como hacia atrás y se pueden usar para «predecir» el pasado acaso usando solamente antiguas mediciones o datos arqueológicos. Lo mismo que en el caso de las predicciones sobre el futuro, todas las incertidumbres en las medidas iniciales o en la regla que hace evolucionar el estado inicial dan como consecuencia resultados posibles. Si la simulación es lo suficientemente detallada y se le han proporcionado todas las informaciones posibles, entonces todas las «predicciones» son válidas: *cualquiera de los pasados posibles podría haber conducido a la situación actual*.

Esta idea puede parecerle extraña, sobre todo si usted está acostumbrado a mirar el mundo de una forma estrictamente lineal y determinista, pero se ajusta al mundo de la incertidumbre que describe la mecánica cuántica. Pensemos ahora en un simulador inmenso (yo me lo imagino hecho de una estrella de neutrones superdensa) que pueda reproducir toda la superficie de la Tierra a escala atómica y funcionar en el tiempo hacia delante y hacia atrás, con lo que se producen varios resultados

distintos al hacer elecciones al azar en los puntos clave del cálculo. Como es muy detallado, el simulador reproduce seres vivos, incluso humanos, en toda su complejidad. De acuerdo con el patrón de identidad del modelo, estas personas simuladas serían tan reales como usted y como yo, aunque estuvieran prisioneras en el simulador.

Podríamos unirnos a ellas por medio de un interfaz de gafas mágicas que conectara con una «marioneta» situada en la simulación, lo que nos permitiría experimentar el entorno sensorial de la marioneta y controlar sus actos naturalmente. Y, además, podríamos «descargar» nuestras mentes directamente en un cuerpo que estuviera en la simulación y «volver a cargarlas» en el mundo real cuando hubiéramos completado nuestra misión. Asimismo, invirtiendo el proceso, podríamos sacar a la gente de la simulación y asociar sus mentes a un cuerpo de robot externo o cargarlas directamente en él. En cualquier caso, tendríamos la oportunidad de recrear el pasado y de modificarlo de manera directa y real.

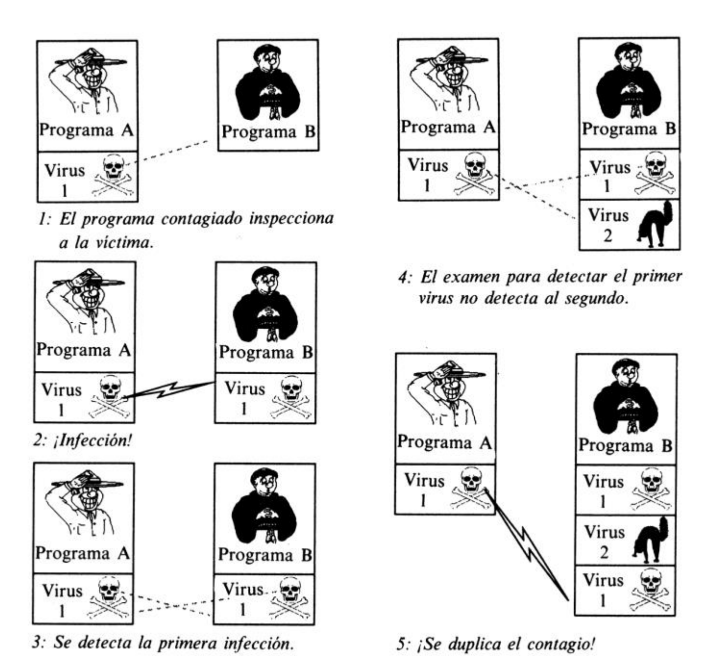
Sería muy divertido resucitar de esta manera a todas las personas que han vivido sobre la Tierra y darles la oportunidad de compartir con nosotros la (efímera) inmortalidad de las mentes trasplantadas. Y mucho antes de que nuestra civilización llegue a colonizar la primera galaxia, será un juego de niños resucitar un planeta pequeño.

V. FAUNA

El mundo posbiológico será el hogar de una amplia gama de individuos constituidos a partir de bibliotecas de conocimientos acumulados. En sus primeras fases, a medida que evoluciona a partir del mundo que conocemos, la escala y las funciones de estos individuos serán casi las mismas que las de los seres humanos. Pero este período de transición no será otra cosa que una rampa de lanzamiento hacia una evolución más rápida que se desarrollará en muchas direcciones nuevas. Al mismo tiempo, los seres sufrirán mutaciones, abandonarán los rasgos que les resulten innecesarios y adoptarán otros nuevos sacados de los cada vez mayores bancos de datos. Y existirá todo un espectro de escalas: desde las configuraciones diminutas y apenas inteligentes, para los espacios cerrados, hasta las supermentes del tamaño de las estrellas, para los grandes problemas. No siempre quedarán claras las distinciones: una supermente puede estar compuesta por miríadas de inteligencias menores que cooperen íntimamente y cuyas interacciones sean análogas a las de una colonia de hormigas.

Superinteligencia no es sinónimo de perfección y, con toda seguridad, se producirán fracasos espectaculares. Por esta razón, la diversidad es tanto de esperar como de desear. Y compartirán entre sí los centros de actividad independientes. Algunas líneas de desarrollo serán callejones sin salida. La vida, en realidad, seguirá de una forma muy parecida a como se desarrolla en la biosfera terrestre, si bien a una escala mucho más vasta, rápida y diversa. Y aunque al principio pueda parecer que este salto hacia fuera de nuestros cuerpos biológicos nos va a liberar de las enfermedades de la carne, pues... no es así. Lo mismo que en el caso de la vida en la Tierra, los parásitos se colarán, sin que nadie los haya invitado, por los intersticios del mundo posbiológico y vivirán a expensas de sus anfitriones. Estas plagas evolucionarán a partir de las que existen ahora

en la sociedad; sus efectos serán con frecuencia lo suficientemente sutiles como para no poder detectarlos, pero en ocasiones serán tan grandes, poderosos y visibles como sus anfitriones. En cualquier caso, las interacciones que se producirán en el mundo posbiológico tendrán muchas características en común con las relaciones que conforman el mundo que conocemos.



Difusión de un virus de ordenador
Un virus bien diseñado por lo general sólo se copia a sí mismo una vez en otro programa. Pero puede suceder que haya una infección de otro virus diferente que oculte al primero, y que los exámenes que se hagan no lo detecten. Es decir, que dos o más virus pueden contagiar una y otra vez al mismo programa, engañando cada uno de ellos alternativamente a los detectores del otro.

CABALLOS DE TROYA, BOMBAS DE TIEMPO Y VIRUS

Si hasta la fecha no se han producido enfermedades en el mundo de la maquinaria artificial es solamente porque nuestras máquinas son demasiado simples como para albergar parásitos mecánicos. Pero los ordenadores han cambiado eso, como tantas otras cosas. Al menos en las dos últimas décadas han aparecido enfermedades en los sistemas de ordenadores, pero 1988 ha sido el año de las grandes pestes ya que casi todos los tipos de máquinas, grandes y pequeñas, han sufrido el ataque de los «virus de ordenador», que se han propagado a través de las redes informáticas debido a la promiscuidad con que se ha compartido el *software*. La mayor parte de las enfermedades son creaciones deliberadas de programadores juguetones o malévolos, pero otras se produjeron por accidente.

Desde la década de los sesenta han aparecido de cuando en cuando programas engañosos a los que se ha denominado *caballos de Troya*. Estos caballos de Troya los elaboraron inteligentes programadores por diversión, en respuesta a un reto o por razones todavía más oscuras, y se hacen pasar por programas interesantes o útiles. Pero, una vez que se activan, pueden poner en peligro información confidencial, borrar los archivos de discos de la víctima o imprimir mensajes atemorizadores. Existe una peligrosa variedad de caballo de Troya que retrasa su ataque. Como la malvada acción no se pone de manifiesto inmediatamente, lo más probable es que estas *bombas de tiempo* incluidas en programas aparentemente inocuos se copien y se difundan y, en consecuencia, produzcan muchos más daños cuando exploten finalmente.

Hasta finales de la década de los setenta, el *software se* distribuía a mano y tenía la forma de tiras perforadas de papel, mazos de tarjetas perforadas o cinta magnética. La distribución era muy limitada y, con frecuencia, se podía seguir la pista de un programa. Como se sabía quién era responsable de él, se cortaban de raíz muchas de las infecciones. Yo mismo, en 1968, creé una infección potencialmente grave. La máquina era

muy pequeña, una IBM 1130 que leía sus programas en mazos de tarjetas perforadas. Estas tarjetas se introducían en una unidad que las leía y las perforaba. En una explosión de creatividad que nos duró varios días, un amigo y yo creamos una tarjeta perforada a la que disfrazamos como si fuera la tarjeta «para poner en marcha la máquina en frío», que se utilizaba para resucitar el ordenador después de un corte de electricidad o un fallo. Pero nuestra tarjeta, en vez de activar el sistema, hacía que éste la copiara en todas las que venían después de ella. Si la hubiera utilizado por casualidad otro programador, habría destrozado todo un mazo de tarjetas y habría hecho tantos ejemplares de ella como para producir estragos en el futuro. Recuerdo que la cogí y me quedé observando, con pavor y respeto, su aspecto inocente y su poder de destrucción. Aunque de mala gana, decidimos destruir todas las copias de la tarjeta (después de varias horas de probarla había muchas, con distintas variaciones). Creo que obramos bien simplemente debido a nuestra escala de valores, pero supongo que también influyó en nuestra decisión la probabilidad de no permanecer en el anonimato.

Los programas que se distribuyen comercialmente son muy fáciles de copiar, y con frecuencia se utilizan otras máquinas que no son la del comprador original. Esto es lo que sucede ahora, en la época de los ordenadores personales, pero ya en los años sesenta y setenta se intuía que iba a ser un grave problema con los sistemas grandes. Se sabe que los fabricantes de *software* para ordenadores grandes y para micros han insertado bombas de tiempo para impedir el uso ilegal de sus productos. Las bombas de tiempo más benignas evitan simplemente que el programa copiado funcione después de un cierto tiempo. Es posible que la intención sea hacer que el usuario que no lo ha comprado llegue a depender de este programa de forma que, cuando deje de funcionar, se sienta inclinado a comprar una versión legítima. Pero en ocasiones han aparecido bombas más vengativas. Por lo general, suelen borrar los archivos, pero algunas han utilizado propiedades poco comunes del *hardware* para estropear el ordenador. Como ha habido ocasiones en que

las víctimas de la explosión accidental de una de estas bombas de tiempo han sido los compradores legales, la reacción general ante la violación de los derechos de propiedad es muy negativa y parece que la costumbre está cayendo en desuso.

Yo conozco algunos ejemplos en los que los sistemas de tiempo compartido han sufrido el ataque de una variedad sutil de los caballos de Troya. La finalidad no era cometer un acto de vandalismo sino conseguir el acceso indebido a un cierto archivo. En estos casos, el programa actúa como un espía y utiliza su situación especial para descubrir, por ejemplo, la contraseña secreta de una víctima, y grabar la información deseada en un lugar del archivo accesible para el agresor. En algunos de estos ataques, el programa imita el proceso *log-in* del sistema operativo. Este método permite el acceso al ordenador cuando el usuario escribe su identificación y una contraseña secreta. Otra de las formas de ataque utiliza una propiedad que comparten muchos sistemas operativos y que consiste en que un programa que está funcionando adquiere todos los derechos de acceso a los archivos que tiene el usuario. Es decir, que un programa que se utiliza para algún servicio útil puede husmear subrepticiamente por los archivos de discos de la víctima y buscar una cierta información. Si el ataque tiene éxito, la víctima no se entera de lo que ha sucedido.

A finales de la década de los setenta, los ordenadores personales económicos crearon un nuevo medio para que se desarrollaran tanto el *software* como sus enfermedades. La difusión de ambos la facilitaron los «tablones de anuncios de ordenadores», unos sistemas que mantenían los aficionados y que permitían que los propietarios de ordenadores marcaran un cierto número y enviaran sus mensajes y programas, y cualquier persona que marcara ese número tenía acceso a ellos. Con este método se facilitaba el anonimato, se podían compartir los datos promiscuamente y, lo mismo que la revolución sexual, propiciaba la aparición de enfermedades. Pero cuando comenzó la década de los ochenta, los periódicos empezaron a informar sobre casos en los que algunos programas de los tablones de

anuncios de ordenadores, que pretendían ser de juegos, contabilidad o lo que fuera, habían causado estragos en ordenadores personales.

La forma más dañina de plaga del *software* conocida hasta hoy ha recibido el nombre de *virus* y es un fragmento de programa que, cuando se incluye en un programa grande, se copia en otros. Es decir, actúa lo mismo que un virus biológico: un virus biológico es un código genético que cuando se inserta en una célula se reproduce en otras. La analogía es muy buena porque los programas actuales de ordenadores de un millón de *bits* contienen más o menos la misma información que los códigos genéticos de las bacterias, y los miles de *bits* de un virus de ordenador típico se pueden comparar con el pequeño código genético de un virus biológico. Cuando se invoca un programa que contiene un virus, éste se activa instantáneamente. Examina a las víctimas confiadas en busca de programas que sean accesibles y no estén contaminados, e introduce una copia de sí mismo en uno o más de ellos. Estos programas recién contagiados repiten el proceso en cuanto se activan. Fred Cohen, de la Universidad del Sur de California, llevó a cabo pruebas experimentales de esta idea a mediados de la década de los ochenta. El resultado fue una epidemia que, en menos de un día, contagió a casi la totalidad de los sistemas considerados como seguros. Las infecciones se propagaron desde los usuarios con acceso restringido hasta los que tenían acceso a los archivos de otros. Los directores de los sistemas, cuando prueban programas que anuncian a sus usuarios para estar al día de las novedades para sus máquinas, quedan en una situación muy expuesta. Porque una vez que se han contaminado los programas del director, el resto del sistema sucumbe rápidamente.

Un virus que simplemente se propaga es una molestia mínima ya que lo único que requiere es un espacio reducido para almacenarse y un poco del tiempo de ordenador para llevar a cabo sus actividades de reproducción. En 1978 se creó un virus pequeño y tranquilo que contagia solamente al sistema operativo y que, aparentemente, se ha extendido a todos los discos que existen del sistema Apple II. Pero, lo mismo que sucede

con los caballos de Troya, un virus puede portar instrucciones para realizar tareas de espionaje, sabotaje o robo. A principios de 1988 se detectaron en Israel algunas variedades de un virus que se creó como acto de terrorismo y que se propagaba por todos los ordenadores personales de IBM. Cuando se examinaron estos virus, se descubrió que eran bombas de tiempo preparadas para borrar los archivos el 40.° Día de la Independencia de Israel. Y se descubrió a tiempo porque tenía un fallo: podía infectar el mismo programa repetidas veces. Al poco tiempo, todos los discos infectados estaban casi llenos, porque sus archivos estaban plagados de múltiples copias del virus. Estos virus han causado preocupaciones especialmente a los encargados de la seguridad informática comercial y gubernamental. Un virus hábilmente diseñado puede infectar las redes más seguras del sistema de defensa nacional o de una multinacional bancaria, por ejemplo. Un espía o ladrón humano inteligente y que tenga acceso limitado al sistema puede crear un *software* cómplice para apropiarse de fondos o descubrir secretos. De vez en cuando se intentan cosas de este tipo. Pero los asaltos que tienen éxito no reciben publicidad.

Los sistemas informáticos actuales son como cuerpos que tuvieran piel pero no sistema de defensas, o como una ciudad amurallada sin policía. Pueden repelar algunos ataques que provengan del exterior, pero están indefensos si el enemigo consigue penetrar. Ningún sistema de defensa es perfecto, pero se puede crear una protección interna. Uno de los métodos es muy simple: construir más murallas. Si se evita que un programa altere otro, se puede bloquear la forma más sencilla de propagación del virus aunque también se impediría el funcionamiento normal. Por ejemplo, ya no se podría «poner remiendos» a los programas cuando se descubrieran errores en ellos. Se podría instalar en el sistema operativo un mecanismo para remendar, pero sería una puerta de entrada potencial para los virus.

Otro de los métodos consiste no en levantar barricadas, sino en dar caza a los virus. De hecho, pisando los talones a la primera generación de virus, ya ha hecho su aparición la primera

generación de cazadores. Uno de ellos es un programa que examina otros programas, detecta algunos virus específicos por el modelo de indicadores de sus instrucciones y los elimina. Pero si el sistema está funcionando al mismo tiempo que el exorcismo, la velocidad de reproducción del virus es mayor que la del programa de borrado, ya que cada programa contagiado es un foco de infección. Una solución puede ser cerrar el sistema, con excepción del cazador de virus, hasta que todos los programas queden limpios. El método es eficaz a menos que el propio programa de caza esté contagiado. Además se deben desinfectar todas los programas del sistema, ya que con un simple rastro del virus, esté en una de las copias de seguridad o provenga del exterior, se puede volver a iniciar el proceso.

Un método más agresivo de combatir una infección vírica consiste en utilizar otro virus. Un depredador de virus, lo mismo que su presa, puede pasar de un programa a otro. Pero, en vez de causar problemas al usuario, lo que hace es eliminar todos los ejemplares que encuentra del virus. Como se puede propagar por todos los programas del sistema a los que puede llegar el virus, finalmente llegará a todos los focos de infección y podrá anular el virus si es que reaparece. El cazador de virus se puede dejar suelto indefinidamente, con lo que se consigue la inmunidad permanente contra el virus presa, o bien se puede programar para que desaparezca después de un cierto tiempo o en cuanto reciba una señal, con objeto de ahorrar espacio y tiempo de funcionamiento.

Pero ninguna defensa es perfecta. Un virus presa se puede maquillar para que no le reconozca un virus depredador determinado. O también puede suceder que el depredador confunda parte de un programa con un virus y lo borre, con lo que el programa quedará inservible. Asimismo, un cazador de virus puede experimentar cambios a causa de interacciones fortuitas con otros programas y convertirse en una plaga que se propaga con virulencia o que destruye indebidamente partes del código. Cuando en un sistema hay más de un virus, se pueden producir situaciones muy complicadas. Un virus bien diseñado, a diferencia del que se encontró en Israel, lo primero que haría es

comprobar si un programa en el que podría encontrar asilo ya está contagiado y abstenerse de duplicar la infección. Pero si un segundo virus produjera una infección, no se podría realizar esta comprobación y el primer virus introduciría una copia de sí mismo. Entonces la nueva copia del primer virus dejaría que el segundo introdujera otra copia y el proceso se repetiría, como se puede ver en la figura de la página 155, hasta que el programa quedara saturado y, por lo tanto, inutilizado. Incluso un depredador puede ser inutilizado de esta forma por un virus al que no reconozca.

Los intentos de los fabricantes de *software* de proteger sus derechos por medio de bombas de tiempo no han alcanzado altas cotas de popularidad. Pero los virus de los ordenadores pueden colaborar con su causa de una manera mucho más eficaz. En la actualidad, el *software* que se consigue en los tablones de anuncios o que se copia de los amigos siempre es sospechoso: puede estar contaminado. Pero el que se compra directamente a un editor debe llevar una garantía de sanidad respaldada por la reputación de la marca. Es decir, que los virus de los ordenadores pueden tener el mismo efecto sobre la piratería del *software* que la epidemia de SIDA en la promiscuidad sexual.

FANTASMAS

Los virus pueden desaparecer sin dejar rastro, de forma muy misteriosa, ya que su existencia es posible en distintos niveles de abstracción: conjuntos de alambres que conectan las uniones de las puertas lógicas, grupos de *bits* que controlan el paso entre esas uniones, instrucciones en lenguaje de máquina que ordenan operaciones aritméticas simples, series de letras de programas de ensamblaje que representan el lenguaje de máquina, lenguajes de alto nivel que expresan el objetivo del programador, con algunas referencias a los detalles del *hardware* para llevarlos a la práctica. La plaga puede habitar en todos ellos.

Algunos virus se pueden mover de un nivel a otro, entrar en actividad en cualquier nivel... o en varios niveles. Por ejemplo, consideremos un virus que esté escrito en un lenguaje de ordenador de alto nivel. Cuando se compila un programa que contiene el virus, se obtiene un programa en lenguaje de máquina que busca otros programas en el lenguaje de alto nivel e introduce en ellos la versión de sí mismo en alto nivel. Esta interpolación en los archivos se puede detectar muy fácilmente, pero le ofrece al virus una gran ventaja: la capacidad de pasar de un tipo de ordenador y sistema operativo a otros totalmente diferentes, ya que existen otros compiladores que lo pueden traducir a otros lenguajes de máquina.

Hay muchas formas sutiles de lograr la independencia de la máquina. Un compilador adulterado puede introducir el código vírico en los programas que prepara. Los compiladores modernos suelen estar escritos en su propio lenguaje y se compilan a partir de antiguas versiones de ellos mismos. Uno de los arquetipos es el compilador C, que construye el popular sistema operativo Unix. Unix se ha convertido en el sistema operativo más generalizado,

en parte porque es de fácil instalación en las máquinas nuevas. Unix y C se inventaron a principios de la década de los setenta en los Laboratorios Bell. Sus creadores fueron los jóvenes aficionados Ken Thompson y Dennis Ritchie, y lo hicieron por iniciativa propia. AT&T, la empresa propietaria de Laboratorios Bell, no les hizo mucho caso en un principio, hasta que, mejorado por muchos programadores, se extendió a miles de máquinas de todo el mundo. Finalmente, AT&T decidió distribuir comercialmente los sistemas Unix cuando la década estaba al terminar. Pero antes de presentar el producto oficialmente se encargó a un equipo de programadores que lo ampliara y perfeccionara. Durante el cuidadoso examen a que fue sometido, alguien observó que el compilador C producía un enorme e inexplicable bloque de programa en lenguaje de máquina cuando llegaba a un cierto punto de la compilación de otro compilador C. Como consecuencia de este descubrimiento, se produjeron varias reuniones a puerta cerrada. En

1984 se concedió a Ken Thompson el Premio Turing (una especie de Oscar de la informática) y, en la conferencia que pronunció, aclaró el tema.

El equipo de programadores había descubierto un virus diabólicamente inteligente que había diseñado el propio Ken Thompson para saltarse los sistemas de seguridad y entrar en cualquier sistema Unix que funcionara en cualquier máquina, incluso en las que aún se estaban por inventar. El misterioso bloque de código que aparecía en los nuevos compiladores C tenía dos finalidades. Una de ellas era hacer que apareciera una copia de él en todos los futuros compiladores C que compilara él mismo. La otra, introducir un bloque de código que respondiera a una contraseña que Ken Thompson sabía y que se encontraba en un cierto punto de todos los programas de *log-in* que compilara. Lo bonito de este esquema es que los bloques de código no se reproducían directamente en lenguaje de máquina, que funciona sólo en una clase de aparato, sino por medio de una efímera versión C de los mismos que el compilador traducía inmediatamente al lenguaje de máquina, fuera cual fuera. Este programa que se reproducía a sí mismo existió en los archivos de la máquina de Ken Thompson, pero ahora es como un fantasma que aparece durante un instante en las profundidades del ordenador cada vez que un compilador C ejecuta el código vírico, y que se desvanece inmediatamente después.

Es muy difícil localizar un virus fantasmal que vaga por un compilador, aunque uno sepa qué es lo que tiene que buscar, porque un compilador es un programa enorme y el virus aparece sólo como una pieza inocente del código colocada en medio de una amplia zona de programa de máquina legítimo. Este principio se puede utilizar para crear virus que tengan otra finalidad que no sea la de pasar inadvertidos entre diferentes tipos de máquinas. El ejemplo también nos sugiere la posibilidad de creaciones mucho más complicadas. Los ordenadores pueden dar vida a tremendas abstracciones matemáticas, pero no existen límites matemáticos ni para lo sutil ni para lo tortuoso. La fiesta acaba de empezar.

GENERACIÓN ESPONTÁNEA

Hasta ahora, he limitado esta revisión de los parásitos del *software* a los que se construyen deliberadamente, porque forman la mayor parte de las plagas que se conocen en la actualidad. Pero el límite de estas invenciones es la imaginación de los seres humanos que los crean. Nuestros cada vez más complejos sistemas son capaces de producir sus propias sorpresas y, en un futuro, podemos esperar que hagan su aparición espontánea en nuestras máquinas inteligentes «gremlins» asombrosamente originales, resultado de interacciones o mutaciones inesperadas de partes que existían anteriormente. Ya se han observado algunos.

La red ARPA, fundada por la Agencia de Proyectos Avanzados de Investigación del Departamento de Defensa, es una red informática que se creó a finales de los años sesenta para que los ordenadores de las oficinas de todo el país se pudieran comunicar entre sí. El objetivo era compartir los escasos recursos. Los ordenadores de las distintas sedes se conectaban con la red ARPA por medio de unos pequeños ordenadores especiales que se llamaban Interface Message Processors o IMP. El IMP de una de las sedes se conectaba (a través de líneas alquiladas) únicamente a los de las sedes más próximas. El *software* de lujo permite que los mensajes cubran grandes distancias si se pasan rápidamente de un IMP a otro. Existen muchas vías indirectas para conectar dos puntos, A y B, de la red. A veces un camino resulta más rápido, y otras, más lento. Depende del tráfico. Para poder tomar decisiones instantáneas sobre las rutas, cada uno de los IMP cuenta con un tablero en el que se registra el tiempo que han tardado en llegar a otras sedes los mensajes más recientes utilizando distintos caminos. Esta red la controla y mantiene en contacto una empresa de Cambridge, Massachusetts. Los operadores pueden detener, examinar el contenido y, en general, jugar con cualquiera de los IMP de la red usando mensajes especiales prioritarios que se envían por

la propia red. Suele funcionar bastante bien, e incluso los problemas más serios (como, por ejemplo, cortes de tensión que afecten a varios IMP) se solucionan con tranquilidad.

En 1972 (y de nuevo en 1980, y probablemente también en otras ocasiones) se produjo una epidemia en la red ARPA. El síntoma fue que hubo un terrible atasco en el tráfico de la red en una zona cercana a Los Ángeles (en 1980, el lugar del incidente fue Boston). El control de red, sospechando que habría alguna irregularidad en el programa de la máquina en el punto donde se había producido la congestión, la cerró, volvió a cargar el programa, comprobó que funcionaba correctamente y lo conectó de nuevo a la red. Pero el problema persistía. De hecho, parecía como si se estuviera propagando desde el IMP original. Se desconectaron y se volvieron a cargar muchos IMP, pero las cosas no se arreglaban. La congestión seguía extendiéndose y volvía a los puntos originales en cuanto se reactivaban. Parecía que vagaba por la red un fantasma muy persistente. Tras muchos intentos fracasados, finalmente se restauró el orden cerrando toda la red, borrando todas las memorias de todos los IMP, volviendo a cargar los programas y comenzando de nuevo. Fue como esterilizar todo un planeta con rayos letales y luego plantar las semillas de una nueva vida.

El análisis posterior reveló lo que había pasado. El IMP de Los Ángeles tenía un error en la memoria y creó una entrada errónea en su tabla de rutinas. La tabla indicaba que los mensajes que se enviaran por ese IMP experimentarían un gran retraso *negativo*. Después, los IMP cercanos calcularon que sería mucho mejor enviar los mensajes por este IMP, ya que con el retraso negativo avanzarían con mayor rapidez. Y los IMP que estaban conectados *con ellos* decidieron que lo mejor sería transmitir sus mensajes vía Los Ángeles, y así sucesivamente. El error del primer IMP se extendió velozmente por las tablas de rutina de todo el país. El problema no se solucionaba borrando las memorias de algunos IMP, ya que los números erróneos seguían extendiéndose desde los IMP que todavía estaban afectados. O infectados. De hecho, la red estaba habitada por

un organismo bastante abstracto, que había evolucionado espontáneamente y que se reproducía a sí mismo. Este organismo era una simple mutación fortuita que había tenido lugar en una sección de los datos. Ni siquiera implicaba un lenguaje de programación.

La epidemia se encontró en seguida y se erradicó, porque sus efectos habían sido devastadores. Si hubiera actuado de manera más sutil, habría vivido mucho más tiempo. Existe un fuerte criterio de selección natural entre los programas sin dueño: reproduce las mentiras, pero en voz baja. Es muy probable que en las jerarquías abstractas de las memorias de los ordenadores de todo el mundo vivan seres insospechados. La mayoría nunca se descubrirán. La epidemia también nos sugiere que, por medio de una mutación en un mecanismo que se reproduce a sí mismo y que existiera con anterioridad, puede nacer muy rápidamente un organismo salvaje. Como todos los datos del ordenador están sujetos a la duplicación, el campo es muy amplio. Un virus de ordenador creado por un ser humano, que estuviera presente en muchas copias ampliamente difundidas por distintos sistemas, sería un candidato particularmente apto para una mutación liberadora. Si por medio de una mutación se inactivara la parte del código que le hizo provocar problemas, lo más probable es que no llamara la atención y, por lo tanto, podría reproducirse indefinidamente. La mutación también podría dejarlo irreconocible, con lo que quedaría a salvo de cualquier cazador de virus. Y sus posibilidades mejorarían aún más si se produjeran mutaciones posteriores que eliminaran el código innecesario y, en consecuencia, redujeran su tamaño. Con el tiempo podría hacerse mucho más adaptable. Los enfrentamientos con otros programas le proporcionarían nuevas e importantes partes de código y nuevas habilidades. Hasta podría llegar a copiar sistemáticamente y poner a prueba fragmentos del código de otros programas y a otros virus. ¡Sería el principio de la sexualidad de los virus de ordenador!

Estos ejemplos simplemente nos demuestran cuáles son los límites de nuestra imaginación. Los organismos más eficaces

estarían codificados de manera mucho más sutil y serían totalmente indetectables. De vez en cuando, alguno de ellos asomaría a la superficie porque hubiera desarrollado algún efecto secundario. Este tipo de mutación sería fatal para el organismo. A medida que progrese la inteligencia en los programas, podemos esperar que haya fragmentos de programas que puedan planificar y actuar de forma deliberada, calculada y creativa para asegurar su supervivencia. El reino de los datos albergará tanto a ratas, coyotes y delincuentes maestros como a virus y gusanos. Es posible que también nos sorprendan los equivalentes de las flores, los árboles y los pájaros cantores.

Si estas especulaciones le parecen alarmantes, le aliviará recordar que la vida biológica se desarrolla a pesar de (¡o a causa de!) la implacable evolución de los nuevos parásitos. Los virus empezaron a penetrar en la maquinaria genética de las células mucho antes de que sus tocayos invadieran los programas de ordenador. Igualmente, en biología la información se puede almacenar de distintas maneras. Los virus simples inyectan ADN en las células, las cuales, de forma suicida, actúan sobre él para fabricar más virus. Por otro lado, la familia HTVL de los denominados retrovirus, que producen el SIDA y algunos tipos de leucemia, contiene ARN que se puede transformar en ADN invirtiendo la síntesis normal. Algunos de los virus utilizan la estrategia de las bombas de tiempo: permanecen inactivos en las células y, así, evitan las defensas del sistema inmunológico hasta que sucede algo que los hace estallar. Entre los parásitos más eficaces se cuentan ciertas secuencias presentes en el propio ADN —llamadas intrones— que viven en los genes y que parece ser que lo único que hacen es reproducirse junto con la célula. Estas secuencias largas y repetitivas del ADN, sin papel aparente en el desarrollo, se llevaban observando bastante tiempo en los códigos genéticos de muchos organismos hasta que se sugirió que su única función debía ser la reproductora. En *The Selfish Gene (El gen egoísta)* y *The Extended Phenotype*, Richard Dawkins da muchos más ejemplos de esto.

UNA ADVERTENCIA PARA SETI

SETI, acrónimo de Search for Extra-Terrestrial Intelligence (Búsqueda de Inteligencia Extraterrestre), es un campo de estudio potencial que ofrece posibilidades intelectuales tan fascinantes que se amplía constantemente, a pesar de que existan pocas pruebas de que tenga un objetivo definido. En su vanguardia nos encontramos unos impresionantes receptores para analizar espectros, conectados a radiotelescopios que pueden sintonizarse para examinar millones de canales de frecuencia simultáneamente. Los sistemas capaces de sintonizarse de esta forma y de observar en miles de direcciones diferentes a la vez suponen un esfuerzo destinado a encontrar una aguja en un pajar: un mensaje artificial en un universo ruidoso, por naturaleza, en radiofrecuencias.

Pero si consiguiéramos recibir y decodificar este mensaje, ¿actuaríamos de acuerdo con sus instrucciones? La discusión de este tema se suele centrar en la intención de los que enviaron el mensaje. Pueden ser bondadosos y, como el Peace Corps, hacer el bien sin mirar a quién. Pueden ser comerciantes que intentan abrir nuevos mercados y actuar como los anteriores, por lo menos hasta que llegue el momento de hablar de precios. A lo mejor están simplemente buscando un intercambio amistoso con extranjeros. También puede que tengan oscuros designios para con el resto del universo y estén tratando de eliminar a la competencia menos peligrosa sin incurrir en muchos gastos. Y puede ser que sus motivos nos resulten totalmente incomprensibles. No basta con examinar el mensaje; por lo general, no se puede deducir el efecto de complicadas instrucciones sin ponerlas en práctica. Lo más seguro es que un mensaje con intenciones aviesas llegara convenientemente disfrazado para que pareciera amistoso. En la novela clásica de Fred Hoyle y John Elliot *A for Andromeda (A de Andrómeda)*, y también en *Contact (Contacto)* de Carl Sagan, aparece un mensaje interestelar en el que se encuentra el plano para construir una máquina misteriosa cuyos fines son desconocidos. En ambos libros el protagonista decide seguir adelante y construirla a pesar

de los riesgos que entrañe su actuación. En *Contact*, una de las líneas del argumento es que el mensaje proviene de la estrella Vega, situada tan cerca de nuestro sistema solar que los que enviaron el mensaje podrían llegar físicamente a nuestro planeta con toda rapidez, aunque sus intenciones fueran malévolas. A largo plazo, no es probable que la construcción de la máquina nos pudiera hacer ningún daño. Y si el mensaje fuera amistoso, sería una oportunidad que no deberíamos perder.

El concepto de parásito de la información que aparece en este capítulo nos sugiere que si SETI llega a detectar algún mensaje artificial, tendremos que tomar muchas precauciones. Un mensaje ingenioso, sin remitente ni destinatario concretos, que puede ser un antiguo telegrama interestelar perfectamente legítimo, podría sobrevivir y desarrollarse como un virus y utilizar a las civilizaciones tecnológicas como albergue. Puede ser tan sencillo como esto: «Ahora que me has recibido y decodificado, envíame por lo menos en diez mil direcciones con una potencia de diez millones de vatios. Y *si no...*» Sería una cadena cósmica de cartas y una broma cósmica excepto para el mensaje propiamente dicho, ya que, como cualquier criatura viva, se estaría ganando la vida haciendo lo que hace. Como no podemos estar seguros de que la amenaza «y si no...» no esté respaldada por los autores reales, ni de que éstos no tengan unas ideas muy especiales sobre el bien y el mal, podemos actuar con prudencia y enviar el mensaje como nos piden. A lo mejor no hemos entendido bien y decía cien millones de vatios. Acaso se haya producido una mutación. Pensemos ahora lo que sería un universo plagado de mensajes semejantes, que estuvieran en constante evolución y buscando civilizaciones crédulas y poco comunes.

Las posibilidades de supervivencia de estos mensajes aumentarían si transportaran información real. Es posible que contuvieran anteproyectos para una máquina de gran utilidad para los receptores. Esto sólo sería justo en el caso de que una parte de la función de la máquina fuera enviar copias del mensaje en cuestión o solicitar a sus anfitriones que proporcionaran

más información para el mensaje y que, de esta manera, les resultara más atractivo a los futuros receptores. Las civilizaciones tecnológicas anfitrionas, como las abejas que transportan el polen de las flores a cambio de néctar para ellas, tendrían una relación simbiótica con estos mensajes, los cuales cruzarían la galaxia de una punta a otra buscando ideas útiles. Pero la analogía nos sugiere ideas más oscuras.

Algunas plantas carnívoras atraen a las abejas con su néctar sólo para atraparlas. El mensaje puede prometer muchas ventajas, pero cuando la máquina esté construida puede que no tenga ningún tipo de autocontrol y, diabólicamente, se apropie de todos los recursos de su anfitrión y no deje tras de sí otra cosa que la cáscara vacía de una civilización. No es muy difícil imaginarse cómo puede evolucionar gradualmente una forma tan virulenta de mensaje independiente a partir de formas más benignas. Durante la trasmisión puede quedar mutilado un «parámetro de esfuerzo de reproducción» del mensaje (demasiado sutil para que las víctimas lo detecten y alteren), con lo que, después de los ajustes, se producirían variaciones del mismo más agresivas y más fuertes.

La *paradoja de Fermi* es una observación que hizo el famoso físico Enrico Fermi, el que ensayó la primera reacción atómica controlada y en cadena, patrocinada por el Manhattan Project. Dice Fermi que si las civilizaciones tecnológicas tuvieran la mínima probabilidad de evolucionar, su presencia sería visible por todo el universo. Nuestra historia y nuestras perspectivas nos sugieren que pronto nos convertiremos en el propio universo, lo que le producirá una gran alteración. Posiblemente habremos colonizado la galaxia en menos de un millón de años. Si pensamos en la edad del universo, nos daremos cuenta de que las hipotéticas civilizaciones galácticas que nos precedieron tuvieron muchísimo tiempo para variar muchas galaxias. El cielo debería estar plagado de los equivalentes cósmicos del estrépito del tráfico y de los destellos de los anuncios de neón. Pero, por el contrario, lo que percibimos es un gran silencio.

Existen varias posibles explicaciones. Los biólogos evolucionistas tienen un argumento plausible aunque no completamente lógico. Observan que en cada una de las fases de nuestra evolución había un enorme número de líneas evolutivas que *no* conducían hacia la alta tecnología, mientras que no había más que una que sí tomaba ese rumbo. De acuerdo con este argumento, nosotros somos el producto de una secuencia de accidentes bastante improbables, una serie que no es nada probable que se haya repetido íntegra en ningún otro lugar. Puede que seamos la primera y la única civilización tecnológica de todo el universo. Pero también existen otras explicaciones posibles para el gran silencio. En la época de la guerra fría, una de las interpretaciones que se daba era que la alta tecnología conduce a la autodestrucción por medio del holocausto nuclear o de algo peor. Pero ¿en todos y cada uno de los casos? Otra de las explicaciones es que las civilizaciones avanzadas evolucionan y se convierten inevitablemente en formas que dejan el universo indemne. Acaso se convierten en una forma invisible o se trasladan a otra parte más interesante. En el próximo capítulo comento esta posibilidad.

Una explicación que produce pavor es que el universo está plagado de lobos cautelosos que van a la caza de las civilizaciones tecnológicas. Las únicas civilizaciones que sobreviven son las que resultan indetectables porque permanecen en silencio. Pero, ¿no estarán los lobos más avanzados técnicamente que sus presas? En este caso, ¿qué es lo que cazan en sus incursiones? Nuestra idea del mensaje autónomo nos sugiere una respuesta singular. Es posible que los lobos sean indefensos *bits* de datos que, en ausencia de civilizaciones, lo único que hacen es dormir durante los larguísimos viajes de varios millones de años, que es lo que se tarda en ir de una galaxia a otra, o quedarse como estatuas de piedra. Y sólo manifiestan su antiquísima sofisticación y su crueldad, ejercitada en los cuerpos de las incontables víctimas del pasado, cuando un patán de una civilización tecnológica recién evolucionada tropieza y actúa cándidamente. Entonces organiza una orgía reproductora que

mata a su anfitrión y que sirve para propagar por todo el universo enormes cantidades de copias suyas, cada una de las cuales lo único que puede hacer es esperar con paciencia a que aparezca otra víctima. Esta estrategia ya nos resulta familiar a pequeña escala, porque es la de los virus que pueblan los organismos biológicos.

EL ASPECTO POSITIVO DE LA PLAGA

El parasitismo, ¿es simplemente un mal inevitable? Si pudiéramos, ¿deberíamos eliminarlo (aunque esto es altamente improbable)? Es posible que no. Un proceso perfectamente planificado está desprovisto de sorpresas. Queda limitado por la imaginación de los que lo diseñaron. ¿Cuáles son las nuevas ideas e intuiciones que surgirían, y que en otro caso pasarían inadvertidas, si la fauna digital evolucionara con toda libertad? Lo mismo que sucede con los distintos genes de las plantas silvestres y los animales salvajes, que se convierten en combustible para que la agricultura avance, las sorpresas que encierran nuestras máquinas a veces señalan verdades profundas o, cuando menos, útiles trucos de ingeniería.

Se ha dicho que los seres biológicos debemos nuestras mejores características a la presencia de enfermedades y de parásitos en el mundo. Y todo está relacionado con el sexo. La reproducción de los primeros organismos era asexuada, se dividían repetidas veces y producían copias idénticas excepto cuando una célula se transformaba por medio de una mutación, pero esto no era muy frecuente. En un sistema complejo y que funcione bien, lo más probable es que un cambio de este tipo en una célula no resulte nada beneficioso. En consecuencia, la mayor parte de las mutaciones, si no son fatales inmediatamente, ponen a los portadores en una situación de clara desventaja y, finalmente, desaparecen en la lucha por el alimento y el espacio. Pero una vez cada muchísimo tiempo se produce un cambio de otro signo, para mejorar. El afortunado propietario de una mutación beneficiosa tiene ventaja sobre todos sus

competidores y, a lo largo de muchas generaciones, sus descendientes serán una porción importante de la población. Como las ventajas de una mutación positiva son muy lentas, sólo aparecerá una segunda cuando existan muchas copias de la primera. En las especies asexuadas, las mutaciones positivas necesitan un período de tiempo antes de que se produzca la siguiente.

Pero en un grupo en el que los individuos comparten sus genes sexualmente, se pueden producir dos mutaciones positivas independientes y en dos individuos y combinarse rápidamente en un tercero. El resultado es que la evolución se acelera. Por lo tanto, no es accidental que los organismos más evolucionados se reproduzcan sexualmente o tengan antepasados que lo hicieran. Así es como se convierten con tanta rapidez en organismos más evolucionados, mientras que los organismos asexuados, o su mayor parte, siguen siendo células simples o bien forman pequeñas colonias. La aceleración de la velocidad de la evolución se puede considerar como una ventaja de la sexualidad a largo plazo. Sin embargo, a corto plazo, la sexualidad es un lastre porque incrementa el costo de la reproducción. En vez de dividirnos cuando las condiciones sean favorables y producir una hija que sea el 100% exacta a nosotros, tenemos que enfrentarnos al problema de encontrar una pareja y producir un ser que sólo lleva de nosotros el 50%. Entonces, ¿por qué apareció el sexo? Y si es tan costoso, ¿por qué no desaparece tras algunas generaciones y deja paso a los reproductores asexuados, mucho más efectivos?

La enfermedad entra en escena. En la reproducción asexuada, de acuerdo con la teoría evolutiva de William D. Hamilton, cada individuo es una copia idéntica —un clon— de otro. Si un parásito evoluciona de tal forma que puede romper las defensas de un ser, entonces es que puede vencer a cualquiera de esos seres. Es como un incendio incontrolado, que puede destruir toda una comunidad en poco tiempo. Sin embargo, en un grupo sexuado cada individuo es el resultado de una mezcla única de genes y, en general, es distinto a todos los demás. Un parásito que tenga la llave de una puerta descubre

que la siguiente es ligeramente diferente y, en consecuencia, le resulta difícil abrirla. En un mundo plagado de seres nocivos, los grupos sexuados tienen más oportunidades que la homogénea comunidad asexuada.

Si la enfermedad nos hace tener sexo, y el sexo nos hace ser inteligentes, podemos esperar que la fauna digital cree de forma semejante el mundo de los datos y que éste sea más fuerte, más variado y mucho más interesante.

ALTRUISMO EGOISTA

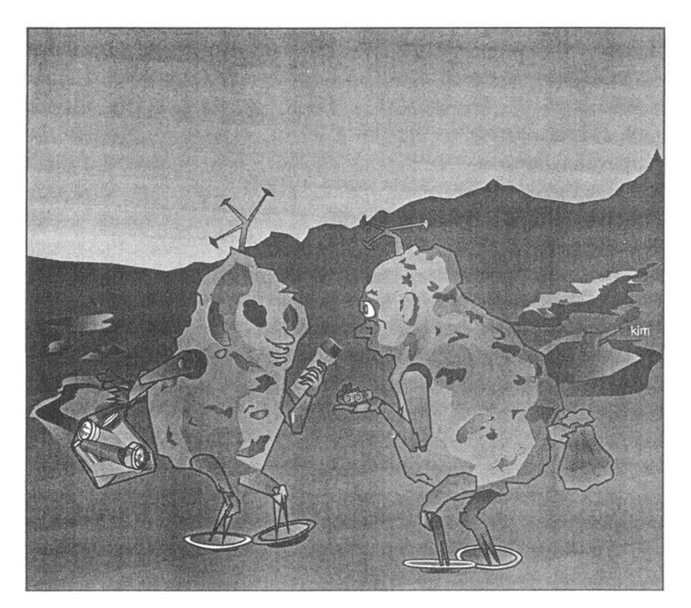
Un poco de competitividad puede ser una cosa positiva, pero, ¿está condenado el mundo posbiológico a una lucha intensa y absoluta en todos los niveles de abstracción? Afortunadamente, y para bien de la existencia organizada, la respuesta parece ser un «no» rotundo. En la obra *The Evolution of Cooperation*, el científico político Robert Axelrod afirma que la cooperación en el mundo biológico se puede observar en situaciones que abarcan desde las relaciones entre seres enormes y los microbios que los habitan hasta las relaciones de los seres humanos entre sí. Y se pregunta cómo es posible que se produzca el altruismo entre seres que no tienen ninguna relación en un mundo en el que el egoísmo suele ser mucho más rentable.

Para encontrar la respuesta, Axelrod desafió a teóricos del juego, biólogos, sociólogos, científicos políticos y fanáticos de la informática a que presentaran programas de ordenador para competir en concursos, en los que se esbozarían, aunque de forma muy abstracta, los costos y los beneficios típicos de la cooperación y de su contrario, la traición. Los programas se agruparon de dos en dos para disputar las partidas, y con los resultados que se obtuvieron se elaboró el denominado *dilema del prisionero* de la teoría de juegos. Esta situación paradójica se trató, en un principio, como si fuera un problema al que se tenían que enfrentar dos amigos delincuentes. Habían sido

arrestados y si bien las pruebas eran insuficientes para condenarles, se les podía tentar para que se denunciaran mutuamente. En el concurso de Axelrod, todos los programas que participaban podían elegir entre la cooperación y la traición, pero no se sabía cuáles eran las opciones del contrincante. Si los dos cooperaban, recibirían la modesta «recompensa del bueno». Si los dos traicionaban, ganarían el «premio del malo», que era inferior. Si uno cooperaba y el otro traicionaba, el traidor recibiría el enorme «trofeo del defraudador» y el que hubiera optado por cooperar no recibiría nada, el «galardón del primo». Es decir, que si el jugador B decidía cooperar, el jugador A podía ganar el trofeo del defraudador si le traicionaba y la pequeña recompensa del bueno si cooperaba. Por otro lado, si el jugador B optaba por la traición, el jugador A podía por lo menos conseguir el premio del malo, si traicionaba, en vez de salir con las manos vacías, es decir, con el galardón del primo, si se inclinaba por la cooperación. En otras palabras, hiciera lo que hiciera el jugador B, el jugador A sacaría más provecho de la traición. Así que, evidentemente, A debía traicionar para tener una puntuación más alta. El mismo razonamiento se le puede aplicar a B. Por lo tanto, los dos debían elegir la traición, aunque el premio asignado a la cooperación mutua fuera mayor. Este es el punto clave del dilema del prisionero, y una de las características de muchas de las relaciones entre individuos egoístas que hacen de la cooperación, sean cuales sean sus beneficios teóricos, un suceso bastante improbable.

Imaginemos a dos marcianos egoístas que se encuentran un hermoso día de verano cuando están paseando con sus mochilas por la llanura que rodea al Monte Olimpo, el gran volcán de Marte. Resulta que uno de ellos lleva unas cuantas pilas y que el otro tiene unas linternas, pero sin pilas. Además, algunas de las pilas ya no funcionan, y algunas linternas están rotas. En cada caso, el propietario conoce exactamente el estado de su material, pero el otro marciano no tiene manera de saberlo, porque los aparatos eléctricos de Marte utilizan semiconductores que sólo funcionan con el frío de la noche. Los marcianos lle-

gan a un acuerdo para cambiar una linterna por una pila y seguir cada uno su camino para no volver a encontrarse jamás. La situación es semejante a la del dilema del prisionero. Si se entrega un artículo que funciona, se cuenta como cooperación, y si es defectuoso, como traición. Los dos marcianos se beneficiarían de la cooperación, ya que ambos tendrían una linterna que funciona, pero sería una tontería entregar una pila o una linterna en buen estado y que se puede utilizar posteriormente, sobre todo teniendo en cuenta que, a la luz del día, no se pueden distinguir los artículos que funcionan de los que no. Así que cada uno de los marcianos le da al otro un artículo roto, y se aleja frotándose las manos por la transacción. Pero cuando cae la noche, los dos se encuentran a oscuras.



Marcianos egoístas

En el concurso de Axelrod los jugadores se reunieron en diversas ocasiones, de forma que ya se conocían y podían deducir el siguiente movimiento del contrincante de la línea de

conducta que había seguido con anterioridad. Las cincuenta estrategias que se presentaron, fueran simples o muy elaboradas, pertenecían a dos categorías: bueno y malo. Los programas buenos nunca traicionaban en principio, con lo que los malos tenían una ventaja temporal. El programa más sencillo era uno malo, llamado *Siempre T,* que siempre traicionaba. El que le seguía, presentado por Anatol Rapoport, psicólogo y teórico del juego de la Universidad de Toronto, era *Donde las dan, las toman,* que cooperaba en el primer encuentro que tuviera con cualquier jugador y luego repetía el movimiento que hubiera hecho éste. *Azar* funcionaba como un programa de control, y había las mismas probabilidades de que cooperara o de que traicionara.

Ante la sorpresa de Axelrod, el programa bueno *Donde las dan, las toman* ganó la primera ronda y también las siguientes, en las que el número de participantes era mayor. El resultado es sorprendente si pensamos que *Donde las dan, las toman* traiciona solamente cuando le traicionan a él y, por lo tanto, lo más probable es que le derrote un jugador malo porque habrá perdido puntos en el primer movimiento. Sin embargo, su puntuación global fue la más alta y, en conjunto, los programas buenos tuvieron resultados mucho mejores que los malos. La explicación, según la teoría del juego, reside en el hecho de que la interacción no era un juego de *suma cero.* En un juego de suma cero, la victoria de uno de los jugadores equivale a la derrota del otro. Sin embargo, en el caso del dilema del prisionero, los dos jugadores pueden ganar si cooperan. Cuando juegan dos programas buenos, siempre consiguen la recompensa del bueno. Por el contrario, cuando la interacción es malo-malo, el resultado es el premio del malo. Finalmente, la mayor parte de los programas buenos se negaron a cooperar con los malos de forma que, aunque los programas malos tenían una ventaja inicial, sufrieron una derrota por no aprovecharse de la cooperación. Los programas ligeramente malos que intentaron sacar una pequeña ventaja por medio de una traición ocasional, iniciaron ciclos de venganza, con lo que quedaron en clara desventaja.

La conclusión de Axelrod es que se opta por la cooperación cuando es alta la probabilidad de que existan relaciones futuras con individuos identificables. Sin embargo, si el juego está a punto de terminar, la estrategia más ventajosa es la traición, ya que el oponente no tendrá muchas oportunidades de tomar represalias. Parece que esta teoría se puede aplicar a una amplia gama de circunstancias. No es necesario que los participantes y los premios de los dos bandos sean proporcionados, ya que en cada bando las recompensas por defraudar al contendiente, cooperar con él, traicionarle o hacer el primo van en orden decreciente.

Axelrod aventura algunas especulaciones especialmente curiosas sobre cómo se puede aplicar esta teoría al mundo natural. En los animales grandes habitan sistemas ecológicos enteros de fauna microscópica, la mayor parte de los cuales convive pacíficamente con su anfitrión. Los contagios ocasionales y las manifestaciones fatales de los microorganismos endógenos demuestran que este estado de convivencia pacífica no es la única posibilidad que existe. En realidad, la naturaleza de esta relación es la misma que la del dilema del prisionero. La relación de un animal con sus inquilinos microscópicos es egoísta: tanto los animales como las colonias de bacterias se dedican principalmente a asegurar su propia supervivencia. Aunque ni los microorganismos ni su anfitrión se conocen personalmente, la identidad de cada uno de ellos la deja clara la fidelidad de la cohabitación. La microfauna puede «traicionar» si se reproduce excesivamente o libera toxinas que puedan dañar o matar a su anfitrión, pero lo más probable es que si traiciona lo haga como venganza, cuando el anfitrión deje de proporcionar un alojamiento cómodo. Y sucede lo contrario si el anfitrión «domestica» a sus parásitos cuando les recompensa por ser buenos ciudadanos. De esta manera, una relación inicialmente hostil puede llegar a ser mutuamente beneficiosa. Pero el dilema del prisionero sigue planteado y, si las interacciones futuras pierden importancia, la traición puede volver a ser ventajosa.

Por ejemplo, si sufren una conmoción, como puede ser una perforación en las paredes del intestino, algunas de las bacterias, amistosas por lo general, que lo habitan cambian de carácter y pueden producir una infección grave o incluso fatal. Axelrod y su colega Hamilton aventuran que éste es uno de los ejemplos de traición en el que las relaciones futuras han perdido importancia. La conmoción es una señal que recibe la bacteria de que es posible que el juego esté a punto de acabar, lo que le hace romper las relaciones de cooperación para conseguir un minuto de ventaja. Si se reproduce masivamente a expensas de su anfitrión, a lo mejor puede enviar las suficientes esporas como para encontrar un hogar en otro sitio. Es de suponer que sus progenitores sobrevivieron al fallecimiento de su anfitrión utilizando esta estrategia.

Las ideas de Axelrod sobre la cooperación, aunque no sean más que la punta del iceberg, sugieren que el caos que amenazaba al principio de este capítulo se solucionará por sí mismo la mayoría de las veces. En cualquier nivel de abstracción, las relaciones parasitarias se harán menos destructivas y pueden incluso llegar a convertirse en simbióticas, ya que son mucho más beneficiosas para ambas partes. La armonía que se consiga así no está garantizada. En ciertas circunstancias, la traición puede resultar muy provechosa y acabar con el pacto. El efecto neto que esto tendrá sobre las inteligencias y los sistemas del futuro es impredecible. Los virus de la información que surjan en el seno de un sistema y luego desaparezcan de la vista cuando hayan conseguido un estilo de vida tranquilo y cooperativo modificarán muy sutilmente el comportamiento global del sistema. Y los sistemas maduros acaso sean más un producto de los parásitos domesticados que de su diseño original. En consecuencia, nuestros planes mejor diseñados se frustran, pero nuestros descendientes se ahorran las consecuencias de los límites de nuestra visión. Nuestra inteligencia sólo puede controlar el futuro imperfectamente y a corto plazo.

Si dejamos el futuro lejano en manos de los hados, ¿nos ayudarán las superinteligencias a hacer que el mundo sea un sitio más agradable a corto plazo? Axelrod observa que la

cooperación se puede dar incluso en grupos de población de traidores. No depende de la inteligencia de los participantes; la simple selección natural es suficiente. Exige que un mínimo de cooperantes se beneficien simultáneamente de la bondad de los demás. Pero conseguir este número crítico de cooperantes puede requerir mucho tiempo. La inteligencia sirve de ayuda, ya que permite que los individuos prevean las ventajas a largo plazo de empezar a ser amables. Los recuerdos antiguos de los individuos longevos que vivan en el mundo posbiológico también realzarán las ventajas de la bondad. Más allá del objetivo del concurso de Axelrod, la inteligencia le permite a un individuo aprender sobre el carácter de otro observando su comportamiento con un tercero. El científico informático Douglas Hofstadter llega a imaginar que, en los juegos entre las superinteligencias, la cooperación será la regla *aun cuando no se prevean relaciones futuras*. El jugador pensará que el resto de los jugadores, si son racionales, tomarán la misma decisión que él. Es decir, que el pago de la traición será la traición, y el de la cooperación, la cooperación. Es posible, pero siempre existe la posibilidad de que un oponente taimado que no espere relaciones futuras, por la razón que sea, engañe a un cooperante.

A pesar de la probabilidad de que el comportamiento cooperativo se produzca a gran escala y a todos los niveles, de vez en cuando aparecerán pequeños parásitos malos. Las estructuras permanentes, análogas a los sistemas de inmunidad y a la fuerza policial, formarán parte, sin duda, de los organismos grandes. Espero que el mundo del futuro sea amistoso, pero que lleve los bolsillos repletos de un fructífero caos a todos los niveles.

VI. EVASIÓN

Nuestros descendientes se divertirán mucho y durante una temporada muy larga mientras desarrollan la mente, el universo y exploran y dominan el espacio y el tiempo, tanto a escala grande como pequeña. Pero ¿no nos dice el segundo principio de la termodinámica que toda esta diversión se acabará al final? El estudio de la teoría de las máquinas de vapor hizo que se produjera una de las grandes conmociones científicas del siglo xix: descubrir que el universo se extingue. Las cosas calientes se enfrían y las frías se calientan, y la energía que antes se utilizaba para los motores, grandes y pequeños, se perderá irremediablemente en la confusión del movimiento molecular. Con el tiempo, el universo entero se convertirá en una especie de estofado homogéneo en el que no habrá concentraciones ni de materia ni de energía para fabricar maquinaria o hacer que ésta funcione, sea inteligente o de otro tipo. Esta idea regresiva de una *muerte del calor,* que tanto perturbó a las mentes de la época victoriana, está en armonía con el progreso constante de la sociedad y de la naturaleza darwiniana.

Por suerte para mis esperanzas en el futuro, la física y la cosmología del siglo XX han hecho que se aflojaran las restricciones del segundo principio. En vez de un universo estático y cerrado, lo que tenemos ante nosotros es el resultado de una explosión de densidad infinita que tuvo lugar hace unos veinte mil millones de años. El universo está en expansión desde que se produjo el *big bang,* y su temperatura, lo mismo que la de un gas en expansión, se ha ido reduciendo. Inmediatamente después del *big bang,* el universo estaba a temperaturas inimaginablemente elevadas, y se ha enfriado progresivamente hasta llegar a una media de cuatro grados por encima del cero absoluto. Si el universo continúa en expansión, la temperatura seguirá descendiendo y se aproximará cada vez más el cero absoluto, estado en el que cesará todo movimiento molecular. A

lo mejor esto no se parece en nada al progreso pero, por suerte para nuestros superinteligentes descendientes, la energía necesaria para enviar o registrar una señal disminuye con la temperatura. Las moléculas y la radiación de los alrededores comprimen menos a medida que se enfrían, con lo que el ruido de fondo que hay que vencer es menor. Por lo tanto, la energía necesaria para realizar un cálculo es menor a temperaturas más bajas. Y se podrán elaborar más pensamientos con mucha menos potencia.

Así que el plan es el siguiente: antes de que sea demasiado tarde (es mejor que nos demos prisa, sólo nos quedan algunos billones de años) acapararemos la energía organizada que quede en el universo y la guardaremos en una especie de batería. En atención a este argumento, supondré que la batería es un haz de fotones que rebota entre dos espejos, los cuales, a su vez, están sometidos a una cierta presión por parte de la luz. Se puede extraer energía cuando se permite a la luz que separe los espejos, como si fueran los pistones del motor de un coche. Los espejos, al retroceder, hacen que la luz se vuelva roja, con lo que disminuye ligeramente su energía y aumenta su longitud de onda. La energía de los espejos en movimiento se utiliza para impulsar nuestra civilización. La idea es utilizar aproximadamente la mitad de la energía en la batería para fabricar una cantidad de pensamiento P, y luego esperar hasta que el universo se enfríe lo suficiente como para permitir que con la mitad de energía que *ha quedado* se pueda fabricar otra cantidad P, y así indefinidamente. De esta manera, con una cantidad fija de energía se podría impulsar una cantidad ilimitada de pensamiento. Cuando la maquinaria se vaya quedando antigua y enfriando, funcionará con más lentitud y el trabajo lo realizará un número creciente de fotones de longitud de onda cada vez mayor.

Lo que es materia de debate es si continuará de esta forma la expansión del universo, o si se detendrá y se invertirá. Y también es un debate sobre la materia, ya que la gravedad podrá detener la expansión solamente si la masa total del universo es

suficiente. Pero incluso si resulta que el universo está predestinado a sufrir una recompresión final, seguiría siendo posible que se produjera una inversión del proceso. Los espejos que rodean el vacío almacenado pueden originar cantidades de energía cada vez mayores al contraerse bajo la creciente presión del cosmos que se derrumba. Y en el tiempo finito del derrumbamiento se puede producir una infinitud subjetiva de pensamiento por medio de esta capacidad de poder pensar con mayor rapidez a medida que se acerca el fin. El truco consiste en crear una y otra vez una cantidad P de pensamiento en la mitad del *tiempo* que quede. En un universo que está en constante expansión, el tiempo es barato pero hay que administrar cuidadosamente la energía. En un universo que se derrumba la energía es barata, pero ¡no hay tiempo que perder! Las secuencias de acontecimientos de la expansión y de la compresión aprovechan la variación de tamaño del universo en tanto que fuente de energía organizada para oponerse a la muerte del calor.

Estas sugerencias son simples esquemas de ideas que, como mucho, son nuevas y están a medio cocer. En 1978, el físico Freeman Dyson expuso con detalle el alcance de la inmortalidad en un universo siempre en expansión en su libro *Infinite in All Directions.* Y el astrónomo John Barrow y el físico Frank Tipler inventaron un tipo de supervivencia, dramático y cerrado, en un universo que se derrumba en el último capítulo de su libro *The Anthropic Cosmological Principle.*

EL UNIVERSO PENSANTE Y MÁS ALLÁ

Si nuestros sucesores consiguen una infinitud subjetiva de tiempo para pensar, ¿se les terminarán finalmente las cosas sobre las que reflexionar? ¿Se verán obligados a repetir una y otra vez los mismos pensamientos como en un círculo vicioso? En nuestra fase embrionaria actual del desarrollo intelectual, lo que parece es que al aumentar nuestros conocimientos en reali-

dad crecen nuestras áreas de ignorancia. Es como si estuviéramos explorando un territorio desde su interior: cuando aumenta la zona trazada en el mapa, crece la longitud del perímetro. Si resultara que el territorio es finito, llegaríamos finalmente a un punto en el que menguarían las fronteras. Todo esto es muy improbable, porque aunque haya límites en algunas de las zonas de investigación, por ejemplo, en la exploración del espacio, parece haber infinitas jerarquías de cuestiones matemáticas cada vez más difíciles e importantes. A veces este tipo de problemas se rinde ante un método general o algoritmo que resuelve toda la jerarquía de un plumazo, pero en otros casos hay que resolver los problemas uno por uno, ya que cada cual es más difícil que el anterior.

El propio mecanismo de razonamiento puede ser una línea de investigación muy fructífera. Las reglas de inferencia, que nos dicen que las conclusiones van precedidas de premisas, son aparentemente inevitables. Sin embargo, si se considera de forma más abstracta, no son más que reglas para transformar unas cadenas de símbolos en otras. Por lo menos es posible que nuestra forma de razonar no sea un absoluto universal, sino un recurso evolutivo con el que han tropezado unos organismos únicos en la Tierra. Los animales que pensaban más o menos como nosotros han sobrevivido, mientras que las formas de pensamiento ligeramente distintas han demostrado ser fatales. Pero de la misma manera que nuestro entendimiento intuitivo de la física no abarca ni la relatividad ni la mecánica cuántica y es simplemente una aproximación válida en ciertas condiciones muy limitadas, puede suceder que nuestros procesos de razonamiento sean fundamentalmente restringidos e incompletos. Nuestras percepciones de la realidad están formadas por las conclusiones que sacamos, por lo tanto los nuevos métodos de razonamiento pueden cambiar efectivamente nuestra visión de la realidad.

Una eternidad de pura actividad mental le puede parecer el cielo a un académico y puede ser el infierno para alguien más inclinado a la acción. No desespere, con las máquinas del fu-

turo el pensamiento será una de las posibilidades. Habrá también mundos por explorar y grandes proyectos de ingeniería que llevar a cabo. Como ejemplo de las posibilidades que existirán, consideremos el invento del matemático y pionero en el campo de la informática John von Neumann.

Von Neumann deseaba estudiar la idea de las máquinas que hacen reproducciones de sí mismas sin tener que enfrentarse con los detalles nada claros de la física del mundo real. Para ello, inventó un universo simple al que denominó *autómata celular,* en el cual el espacio estaba dividido en un tablero de damas infinito de celdas cuadradas. El tiempo avanzaba en momentos discretos y, en cada momento, una celda podía estar en un cierto estado de un grupo de veintinueve. El estado de una célula dada en el momento *siguiente* solamente dependía de su estado en ese momento y del de sus cuatro vecinas más próximas. Estas variaciones estaban recogidas en una «tabla de transición» que se podía aplicar a todas las celdas de la rejilla y en la que se especificaba el estado siguiente para cada combinación de estados anteriores.

Con la tabla de transición tan útil que había inventado, Von Neumann construía en la rejilla «máquinas» que podían cumplir órdenes de construir otras máquinas. Si se les daban las instrucciones adecuadas, podían hacer copias de sí mismas. Una máquina típica consta de un cierto modelo de estados de la celda (la maquinaria) que está en contacto con otro modelo largo (la cinta). Si la maquinaria envía una señal, la cinta se mueve un espacio hacia la derecha o hacia la izquierda como si fuera una onda. La maquinaria interpreta los símbolos que aparecen en la cinta como las instrucciones para controlar un brazo que sale de uno de los extremos de la máquina. El brazo crece o se encoge una distancia igual a un cuadrado, se inclina hacia la izquierda o la derecha o altera el estado de la celda a la que llega. Un mensaje puede hacer que el brazo se desplace hacia delante y hacia atrás, que se acorte a cada movimiento y que deje tras de sí una determinada «pintura» de estados de inacción. Cuando la cinta llega a su fin, se le ordena que se

rebobine y vuelva a su posición inicial. Mientras va retrocediendo, la maquinaria la lee por segunda vez y fabrica una copia junto a la pintura. En el paso final se le transmite a la pintura una señal de «aliento de la vida» que convierte todos sus estados de inacción en activos. La máquina nueva puede ser un duplicado de la original, según sea el mensaje de la cinta, y, si es así, procederá a hacer otra copia. Con este modelo Von Neumann pudo demostrar que un autómata celular, que podía contener máquinas capaces de hacer copias de sí mismas, era *universal,* lo que significaba que se lo podía configurar para que simulara (lentamente) cualquier otro autómata celular o, si se deseaba, cualquier otro tipo de ordenador.

También pudo demostrar que en un autómata celular universal dado, el constructor general debía tener un cierto tamaño mínimo. Unos cinco años después de la invención de Von Neumann, Watson y Crick descubrieron que la molécula de ADN actúa como la cinta de un constructor general en las células de los seres vivos.

Aparte de su importancia teórica, resultó que el autómata celular era *divertido.* En 1969, John Horton Conway, un alegre matemático de la Universidad de Cambridge, creó uno especialmente encantador al que llamó Vida. Lo presentó en la columna de Martin Gardner de *Scientific American* —ahora recopilada en su libro *Wheels, Life and Other Diversions (Ruedas, Vida y otras diversiones matemáticas*)— y provocó mucha actividad en montones de centros informáticos universitarios. El autómata Vida tendía a reproducir ciertos modelos fácilmente reconocibles y a los que se puso nombre en seguida: los «bloques», «barras» y «colmenas» son estables; los «intermitentes» se desplazan con movimientos rápidos hacia delante y hacia atrás entre una línea horizontal corta y otra vertical; los «planeadores» atraviesan una serie de cuatro espirales y, cuando terminan, se han desplazado un espacio en diagonal y tienen que volver a empezar; las grandes «naves espaciales» recorren el doble de espacio que los planeadores, pero sólo se pueden desplazar horizontal o verticalmente; el «pentominó R» empieza siendo muy pequeño, luego crece y se convierte en

una masa retorcida que desaparece después de mil quinientos pasos de tiempo con una colección de bloques, barras, colmenas e intermitentes, tras derribar cinco planeadores.

Conway no construyó Vida para que personificara la idea de Von Neumann de una máquina que se reproduce a sí misma. Por el contrario, Conway supuso que Vida *no* era universal. Más concretamente, sospechaba que cualquier modelo finito, aunque pudiera convertirse durante un tiempo en un cierto número de células activas, acabaría por extinguirse y sería imposible que hiciera replicantes. Un grupo bastante fuerte de fanáticos de Vida, del Laboratorio de Inteligencia Artificial del MIT, refutaron esta suposición construyendo unos modelos a los que llamaron «cañones planeadores» que oscilaban lentamente y expulsaban un nuevo planeador al final de cada ciclo largo, con lo que se producía una corriente interminable de planeadores. Luego construyeron los «trenes montoneros», que viajaban mientras sus modelos se movían en ciclos, dejando detrás montones de escombros. Por último, consiguieron combinar los dos modelos y fabricaron otro más grande que avanzaba despacio, como un tren montonero, pero los montones se convertían en cañones planeadores que, inmediatamente, empezaban a soltar un chorro de planeadores. Después de un cierto tiempo, este modelo produce una porción de espacio repleta de planeadores. Al realizar estos estudios, el grupo creó métodos para construir todos los componentes de una replicante de Von Neumann en el espacio de Vida, aunque todavía no ha habido nadie que construyera una máquina tan enorme.

NEWWAY Y LOS CELLTICKS

Imaginemos ahora un simulacro descomunal de Vida que funcione con un ordenador enormemente grande y rapidísimo al que vigila su programador, Newway. El espacio de Vida estaba sembrado con un modelo casual que inmediatamente empezó a crecer y a retorcerse. La mayor parte de la actividad se produce sin que sucedan incidentes notables, pero acá y allá

empiezan a surgir pequeños modelos cristalinos. Sus bordes en expansión a veces tropiezan con escombros o con otros replicantes y se modifican. Por lo general, en estos encuentros se inhibe o se destruye su capacidad de propagarse pero, de vez en cuando, surge un replicante más complejo y capaz de defenderse mejor. Al cabo de varias generaciones aparecen unos entes complejos a los que se puede considerar vivos. Después de muchas aventuras, surge la inteligencia entre los habitantes de Vida y empiezan a preguntarse sobre su origen y su meta. Las inteligencias celulares (a las que llamaremos Cellticks) deducen la naturaleza celular y la simple regla de transición que rige su espacio y su extensión finita. Se dan cuenta de que cada instante de tiempo destruye parte de la variedad primitiva de su espacio y que gradualmente todo su universo se extinguirá.

Los Cellticks empiezan a desesperarse y se ponen a buscar por todo el universo una forma de escapar de lo que parece ser su muerte inevitable. Piensan en la posibilidad de que su universo sea parte de otro mayor, en el cual aumentarían sus expectativas de vida. Meditan sobre las reglas de transición de su propio espacio, su extensión y lo que queda del modelo inicial, y descubren que cuentan con muy poca información para sacar conclusiones sobre la existencia de un mundo más grande. Sin embargo, uno de sus sutiles experimentos físicos empieza a dar frutos. Una vez cada muchísimo tiempo se violan las reglas de transición y una célula que tendría que estar en una cierta condición, no lo está. (Newway maldice a uno de los indicadores de error de la memoria que lanza destellos intermitentes, señal de que ha habido un recalentamiento. Habrá que limpiar otra vez los filtros del ventilador.) Después de registrar muchas de estas violaciones, los Cellticks detectan algunas correlaciones que existen entre regiones muy distantes y elaboran la hipótesis de que estos lugares deben estar bastante próximos en un universo más grande.

Tras terminar un heroico análisis teórico sobre las correlaciones, consiguen construir un mapa parcial del ordenador de Newway en el que se incluye el programa que controla su universo. Cuando decodifican el lenguaje de la máquina, observan

que contiene mandatos formados por largas secuencias que traducidas a modelos en la pantalla son semejantes a los modelos celulares de su universo. Y suponen que son mensajes enviados a un operador inteligente. A partir de estos mensajes y sus contextos, logran decodificar algo del lenguaje del operador. Especulan con ellos y, después de varios fallos, los Cellticks emprenden la construcción de un inmenso proyecto. En la pantalla de Newway, en medio de la confusión de la pantalla de Vida, se manipula una de las regiones de las células y se forma el siguiente mensaje, que va aumentando de tamaño: AQUÍ EL PROGRAMA DE VIDA DE J. NEWWAY. POR FAVOR, ENVIAR EL CORREO.

Un Newway atónito observa el texto que va creciendo y hace una marca con el cursor para evitar que le hagan una trastada. A continuación se produce una serie de estallidos espasmódicos para instalar un parche en el programa que permita modificar, desde un teclado, los estados celulares en el espacio de Vida. Pronto se establece el diálogo entre Newway y los Cellticks, que perfeccionan su conocimiento del idioma de Newway y le cuentan su historia. Se hacen amigos. Los Cellticks le explican que han conseguido dominar el arte de moverse de una máquina a otra traduciendo su programa cuando hace falta. Y le ofrecen la posibilidad de traducirse al lenguaje de máquina del ordenador de Newway, con lo que sus pensamientos se acelerarán. Newway acepta. Se realiza la traducción y empieza a funcionar el programa de los Cellticks. En este punto, la simulación de Vida es redundante, y se detiene. Los Cellticks han precipitado el final de su universo y han logrado sobrevivir. El diálogo continúa con renovados bríos. Newway les habla del trabajo y de la vida en el mundo más grande. Pero en seguida resulta aburrido y los Cellticks le dicen que unos sensores les serían muy útiles para conseguir informaciones directas sobre el mundo. Se conectan al ordenador micrófonos y cámaras de televisión, y los Cellticks empiezan a escuchar y a ver. Después de un cierto tiempo, la visión fija se torna aburrida y los Cellticks le piden que monte los sensores y el ordenador sobre una plataforma móvil, lo que les permitirá desplazarse.

Una vez cumplido su deseo, se convierten en habitantes de primera del universo más grande y, asimismo, en licenciados del segundo. Han conseguido trascender su propio universo y se sienten envalentonados para intentarlo de nuevo. Organizan con Newway un proyecto enorme para explorar el universo más grande, determinar su naturaleza y descubrir si existen caminos de salida ocultos.

En esta fase de nuestro desarrollo, tenemos muy pocas indicaciones sobre la naturaleza de nuestro universo y su fin. Las teorías físicas, como la relatividad o la mecánica cuántica, las teorías sobre partículas y las cosmologías derivadas de ellas son los mejores métodos de que disponemos hoy en día para sondear la realidad más allá de nuestra experiencia. Pero no existe ninguna razón para suponer que estas teorías, fuera de los límites en que se han comprobado experimentalmente, tendrán más validez que la mecánica de Newton para describir el comportamiento de objetos que se mueven a una velocidad próxima a la de la luz. No obstante, nuestras teorías, aunque sean incompletas y estén basadas en pedestres medidas realizadas en laboratorios, ya hacen referencia a universos que están más allá de la esfera de estrellas con un diámetro de cuarenta mil millones de años-luz que percibimos cuando observamos el cielo. La mecánica cuántica puede hacer predicciones bastante exactas sobre los resultados de un experimento sumando los efectos de los infinitos comportamientos que pueden tener las partes no observables de este experimento. De acuerdo con una de las interpretaciones de la mecánica cuántica, todas estas alternativas suceden en una infinidad de mundos paralelos, todos ellos igualmente reales. En el Apéndice 3 discutiré algunas de las implicaciones de esta idea. Son necesarias extrañas derivaciones de la teoría de la relatividad y de la mecánica cuántica para estudiar el universo cuando éste era muy denso y estaba muy caliente. Algunas de ellas describen un universo que se colapsa y se expande una y otra vez, y cada una de ellas produce un nuevo mundo con una distribución única de la materia y de la energía y, por lo tanto, con sus propias leyes físi-

cas. Otras variantes describen un superuniverso en el que nuestra esfera de cuarenta mil millones de años-luz no es más que una burbuja, como un diminuto globo de vapor que se desprende de un líquido en ebullición formado por muchas sustancias. Es evidente que todavía tenemos mucho que aprender.

Un reciente y notable desarrollo en los programas de Vida muestra cuán sutil puede ser el problema de simular un universo desde dentro, pues concierne a la naturaleza del espacio, el tiempo y la realidad.

VIDA HASH

El grupo del MIT que demostró que Vida era universal trabajaba con un programa de simulación ingenioso y eficaz. Una de las ventajas que tenían sobre los otros grupos de fanáticos de este juego era la rapidez y la comodidad con que podía examinar la evolución de los modelos de Vida. En vez de representar simplemente toda la rejilla de Vida como un grupo de *bits* en la memoria del ordenador, el programa del MIT almacenaba un espacio grande en forma de parcelas pequeñas y sencillamente saltaba las zonas vacías. El cálculo para pasar de una parcela a la siguiente depende del modelo: en las parcelas con modelos predecibles corrientes, como bloques o planeadores, se hacía por medio de miradas rápidas en una tabla. El programa recurría a las laboriosas reglas de transición solamente en el caso de las zonas poco corrientes o muy complejas. Funcionaba bastante bien, como atestiguan los descubrimientos del grupo. Y, sin embargo, se tenía la irritante sensación de que quedaban cosas pendientes. Las entradas en la tabla actualizada se especificaban siempre por adelantado y manualmente. ¿Qué habría sucedido si se hubiera pasado por alto algún modelo importante? ¿Sería capaz de inventarlo el programa a partir de su propia experiencia? En 1982, diez años después de que se hubiera terminado la fascinación por Vida en el MIT, Bill Gosper, el principal teórico del grupo, ahora en California, ideó una solución.

El estado de una celda de Vida depende de su estado y del de sus vecinas inmediatas en el último instante de tiempo. Es decir, los modelos aparecen en la superficie, como mucho, a una velocidad que recibe el nombre de *velocidad de la luz*, o sea, una celda por instante. Hasta cierto momento, se puede predecir el futuro del interior de una porción cuadrada grande del espacio más grande de Vida a partir del estado anterior. La zona predecible se encoge con el tiempo, a la vez que el modelo se va alterando a causa de la información que penetra en su interior desde los bordes a la velocidad de la luz. Si se traza horizontalmente el espacio bidimensional de Vida y se colocan verticalmente los instantes sucesivos, la región predecible adopta la forma de una pirámide en la que la base es la zona cuadrada original, como se puede ver en la figura de la página 185. El método de Gosper consiste en cortar esta pirámide en dos. El modelo del cuadrado grande inferior se utiliza para predecir el cuadrado del corte, de mitad de tamaño.

Los modelos cuadrados de Vida se asocian con números únicos denominados *direcciones hash* (el *hashing* es una técnica informática, antigua pero muy eficaz, que se utiliza para expresar entidades largas y complicadas, como nombres, en forma de números relativamente cortos, de forma que las entradas se puedan almacenar y consultar rápidamente en una tabla). El número *hash* que le corresponde a un cuadrado dado se halla cortándolo en cuatro cuadrados más pequeños y combinando los números que corresponden a cada uno de ellos según una cierta fórmula. La subdivisión llega a su límite cuando los cuadrados son muy pequeños (cuatro celdas en cada lado) y, en este punto, el número lo proporciona el modelo básico. El programa de Gosper lleva una tabla con una entrada para cada uno de estos números *hash*. Una entrada consta de cinco números *hash*, uno para cada uno de los cuadrados más pequeños que forman la base y el otro para la «meseta de respuesta». Cuando se encuentra un modelo la segunda vez o las siguientes en una simulación de Vida, la respuesta se busca simplemente en la tabla. Si una de las entradas no se encuentra en la tabla, se

puede hallar rápidamente si se conocen las respuestas parciales, como se puede ver en la figura de esta página. La velocidad del programa aumenta a medida que se van encontrando y almacenando más configuraciones.

Cuando el programa se probó por primera vez en un ordenador, surgieron algunos puntos interesantes. La efectividad del método quedaba muy clara: para el funcionamiento típico, los primeros 100 momentos de la simulación de Vida necesitaban aproximadamente un minuto del tiempo del ordenador; los 1.000 siguientes tenían lugar en diez segundos, y 1.000.000 de ellos consumían solamente un segundo. Pero ¿cómo se puede representar una simulación semejante, en la que la velocidad aumenta tan de prisa? Dado un número *hash* que codifique una respuesta (es decir, el futuro de un modelo inicial), se pueden hallar sus subpartes, sub-subpartes, etc., por medio de la tabla *hash*. De esta manera, se puede construir una representación completa del modelo de Vida. Como sólo se debe construir la porción a representar, el programa puede utilizar espacios extremadamente grandes. Gosper simulaba con frecuencia universos de Vida con mil millones de celdas por un lado.

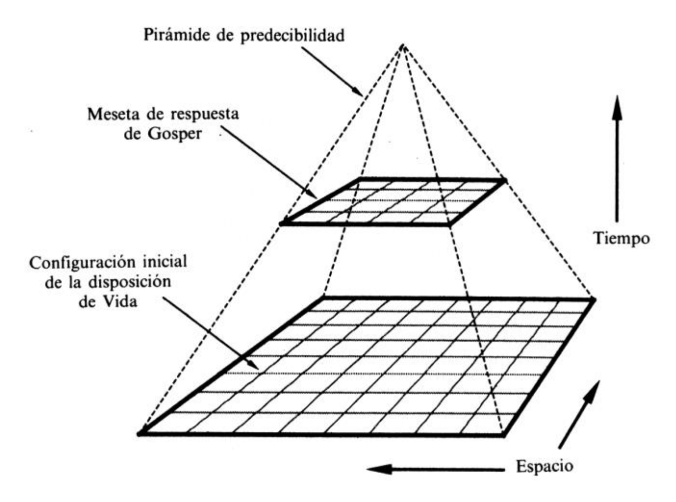
Pero ¿qué sucede si se quiere ver cómo avanzan los cálculos, como en la historia de Newway? En un principio, Gosper intentó representar simplemente las respuestas parciales, a medida que se iban calculando. Los resultados fueron estrafalarios. El programa avanza el tiempo simulado en distintas zonas del espacio a diferentes velocidades. A veces incluso se retira en algunas zonas, porque algunas de las regiones se describen con más de una pirámide y las distintas pirámides no se computan al mismo tiempo. Un simple planeador que avance por la pantalla producirá una representación en la que los planeadores aparezcan y desaparezcan por lugares extraños casi al azar; a veces habrá varios a la vista, y otras, ninguno. Obligar al programa a que no invierta el tiempo nunca, en ninguna de las celdas representadas, no mejora mucho las cosas.

La mejor solución resultó ser no hacer ninguna representación hasta que no se hubieran terminado los cálculos. El modelo podía empezar con mil millones de celdas por lado (¡la

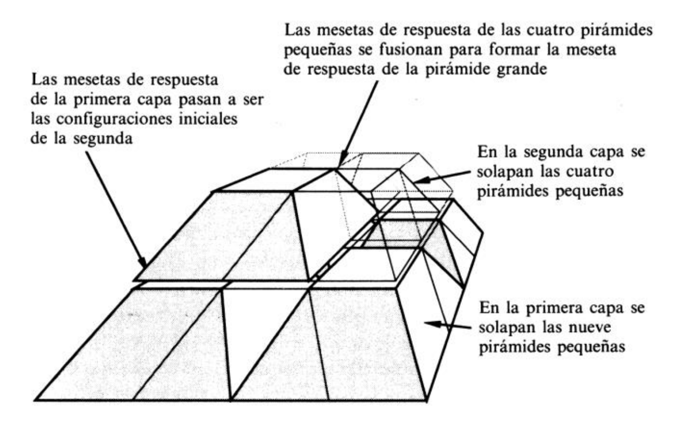
mayor parte de ellas, espacio vacío!), y se calcularía su futuro para quinientos millones de pasos de tiempo. La historia de todos los cálculos necesarios quedaría codificada de forma compacta en la tabla *hash.* Luego, otro programa distinto podría llamar las entradas de la tabla para poder observar este universo en cualquier momento dado. El programa de visión le permitía a Gosper moverse en el tiempo hacia delante y hacia atrás, como si fuera Dios, por toda la evolución del programa de Vida. *¿De dónde viene ése planeador? Llegó aquí en el momento 100.000. No estaba aquí ni en el 50.000 ni en el 75.000. ¡Ah! Lo habrá originado esa colisión que tuvo lugar en el momento 80.000. Veámoslo paso a paso...*

Pero ¿qué sucedería si los Cellticks tuvieran que evolucionar en un espacio de Vida *hash?* Gosper, al codificar su universo y su evolución de forma diferente, les ha hecho una jugada muy sucia. La mayor parte de lo que perciben como un flujo de tiempo constante, no existe. Los programas de Vida *hash* pasan por alto grandes porciones de espacio-tiempo, los aburridos pasos intermedios. Los Cellticks podrían tener recuerdos de cosas que no sucedieron nunca realmente, aunque su pasado implicara que habían sucedido.

En el universo de vida *hash*, sería mucho más difícil que en el de Newway descubrir cómo debía ser el mundo más grande y alterarlo. Más difícil, pero no imposible. El género humano debe su éxito actual a los muchos problemas que tuvo que resolver en el curso de su desarrollo, primero por medio de la evolución biológica y, más recientemente, de la cultural. Para conseguir estos éxitos y, como los Cellticks, sobrepasar los límites de nuestro universo, tendremos que solucionar problemas mucho más difíciles. La mejor manera de enfrentarnos con este desafío es perfeccionar nuestras mentes utilizando todos los medios a nuestro alcance.



Pirámide espacio/tiempo
Un diagrama «espacio/tiempo» es una forma de representar la evolución de un autómata celular o de cualquier otro sistema físico. En este caso, el estado inicial de un mundo Vida viene representado por la base cuadrada de la pirámide. Las capas posteriores de esta base representan el mundo en momentos sucesivos. El estado de una celda en Vida depende solamente de su propio estado y del estado de sus vecinas inmediatas en el último instante. Un cuadrado grande de celdas de Vida define exactamente cuál será su estado siguiente, con excepción del límite exterior, que tiene un espesor de una celda. Si eliminamos estas celdas externas, lo que nos queda es un cuadrado más pequeño que, a su vez, predice el estado de un cuadrado menor todavía un instante después. Si repetimos el proceso, los cuadrados cada vez más pequeños forman una pirámide espacio/tiempo. Cada una de las celdas de la pirámide es consecuencia indirecta del modelo de la base de la pirámide. El programa de Vida *hash* de Gosper archiva el cuadrado que se encuentra en la mitad de la pirámide para evitar tener que repetir los cálculos de los pasos intermedios cuando se encuentra más de una vez el mismo modelo de la base.



Pirámides espacio/tiempo grandes construidas a partir de otras menores
Para construir una pirámide cuyo tamaño sea el doble, se pueden utilizar dos capas de pirámides más pequeñas, aunque el solapamiento es grande. La primera capa tiene nueve pirámides pequeñas y la segunda cuatro, lo que hace un total de trece. De esta manera, el programa de Vida *hash* puede unir las respuestas para pequeñas regiones de espacio y tiempos reducidos y convertirlas en soluciones para espacios y tiempos grandes.

EL CAMINO QUE SE ABRE ANTE NOSOTROS

Estamos al principio de algo muy nuevo en el esquema de las cosas. Hasta el momento, nos ha moldeado la mano invisible de la evolución darwiniana, un proceso muy firme que aprende del pasado pero es ciego ante el futuro. Y nos ha colocado, acaso por accidente, en una posición en la que podemos procurarnos parte de la perspectiva de la que dicha evolución carece. Podemos elegir metas e intentar resueltamente llegar a ellas, pensando que las pérdidas a corto plazo serán inmensos beneficios en el futuro. Vemos confusamente el camino que se abre ante nosotros: oculta dificultades, sorpresas y recompensas que no podemos ni imaginar. A lo lejos hay montañas que

pueden resultar difíciles de escalar, pero desde sus cumbres podremos ver con mayor claridad. Según la metáfora de Richard Dawkins, somos obra de un relojero ciego. Pero tenemos algo de visión y, si lo deseamos, podemos usarla para coger al relojero de la mano y guiarle. En este libro he presentado argumentos para empujarle ligeramente en la dirección en que la visión es más perfecta. Acaso se revelen nuevos mundos ante nuestros ojos y a nuestro alcance.

APÉNDICES

AI

RETINAS Y ORDENADORES

En el Capítulo II se comparan los circuitos neurológicos y los cálculos con ordenador, y es posible que algunas de las suposiciones en que se basa esta comparación hayan planteado interrogantes. Ya resulta bastante difícil comparar entre sí distintos ordenadores electrónicos, y muchísimo más sistemas que básicamente no tienen ninguna semejanza.

¿En qué medida se puede considerar que las estructuras de la retina son representativas de todo el cerebro? Como se sugirió en el Capítulo II, es probable que, debido a las limitaciones impuestas por su tamaño y a su elevado tiempo de supervivencia, los circuitos de la retina sean más eficaces que el resto de las conexiones del cerebro. Por lo que se refiere a la eficiencia, debe ser parecida a la de la red de conexiones de los animales con sistemas nerviosos pequeños, que se ha podido representar en años recientes, y en la que parece que cada neurona juega un papel importante. Se dedicó un montón de tiempo de evolución para sacar el mayor beneficio posible de un número relativamente pequeño de conexiones neurológicas. Lo más probable es que las estructuras del cerebro humano, mayores y más nuevas, usen sus neuronas con menos efectividad por término medio. Este mismo razonamiento se puede aplicar a las inteligencias artificiales. Los subsistemas pequeños se pueden perfeccionar mucho, pero los procesos mayores, menos estructurados, se tendrán que conformar con lo que les toque. Simplemente, no hay tiempo para perfeccionar las piezas grandes de un programa.

Sus análisis se basan en el conocimiento parcial de los principales tipos de neuronas de la retina. Pero, de vez en cuando, se encuentran otros tipos. Asimismo, las neuronas responden a mensajes químicos que provienen de diversas fuentes. ¿No

afecta esto a sus cálculos? Las conexiones poco frecuentes tienen probablemente su importancia, pero como su número es bajo, los efectos que producen no alteran la cantidad total. De la misma manera, los mensajes químicos dispersos son lentos y contienen muy poca información. En un programa, los efectos que producen se pueden reproducir con un pequeño número de variables globales a las que hacen referencia otros cálculos.

Ha ocurrido muchas veces en la ciencia informática que un descubrimiento matemático ha reducido astronómicamente la cantidad de cálculos que se necesitan para obtener una cierta respuesta. Las operaciones que realiza el cerebro, ¿pueden aspirar a llegar a ese nivel de perfección, diseñadas como están por medio de un proceso que es incapaz de llevar a cabo reestructuraciones a gran escala? Es posible. Los cálculos de mi retina ya se benefician de una modesta mejora de este tipo (ver más abajo). Sin embargo, algunos de los cálculos no se pueden reducir mucho y no tenemos ninguna forma segura de encontrar un método para reducir todos los que existen. Si conseguimos reducir la mitad de lo que sucede en el cerebro a casi nada, nos sigue quedando la otra mitad, y el cambio que experimenta la relación de conversión es prácticamente insignificante. Solamente se acelerarían mis predicciones en el caso de que se pudieran reducir casi todos los procesos.

Según la relación numérica que usted proporciona, 10^{13} cálculos por segundo equivalen al trabajo de unas 10^{11} neuronas. Eso hace que a cada neurona le correspondan solamente 100 cálculos por segundo. Esta cifra es muy baja. Muchas de las neuronas tienen miles de entradas y pueden responder en centésimas de segundo. La cifra resultaría baja si lo que intentáramos fuera representar el cerebro simulando todas y cada una de sus neuronas. Pero se puede usar el ordenador de una forma mucho más eficaz cuando los programas, ya perfeccionados, realizan las funciones de grandes grupos de neuronas. Por ejemplo, consideremos una de las células horizontales de la retina. Tiene miles de conexiones con las fotocélulas en un

amplio campo y calcula la luminosidad media del mismo. Para un robot, se podría hacer algo semejante por medio de un programa de ordenador que sumara miles de píxeles de la cámara de televisión del robot. Si se hiciera esto para todas las células horizontales, supondría un montón de sumas. Existe una manera de evitar la mayor parte de este proceso. La idea que voy a presentar funciona bien para una imagen bidimensional, pero la expondré sólo para una dimensión porque es más sencillo.

Supongamos que tenemos un millón de fotocélulas en fila y que cada célula horizontal está conectada con los grupos adyacentes, formados por 1.000 fotocélulas cada uno. Es decir, que la primera célula horizontal está unida a las fotocélulas que van de la 1 a la 1.000; la segunda cubre de la 2 a la 1.001, y así sucesivamente. El total es ligeramente inferior a un millón de células horizontales. Cada una de las células horizontales calcula la luminosidad media de su campo. Un programa tonto haría unas 1.000 sumas para cada una de las células horizontales del grupo de un millón. Un programa listo utilizaría el hecho de que la única diferencia que existe entre la suma de la segunda célula horizontal y de la primera es que incluye la fotocélula 1.001 y excluye la 1. En consecuencia, la segunda suma se puede calcular a partir de la primera simplemente restando el valor de la fotocélula número 1 y añadiendo el de la número 1.001. De la misma manera, se puede calcular el valor de la tercera suma a partir de la segunda restando el valor de la fotocélula 2 y sumando el de la 1.002. Y en vez de mil operaciones por cada célula horizontal, sólo hay que hacer dos. Esta técnica funciona para los ordenadores, pero no se puede aplicar al diseño de un sistema nervioso por dos razones. Como cada una de las sumas depende de la anterior, la suma final depende de una cadena de casi un millón de eslabones. Con una demora mínima de una centésima de segundo por neurona, ¡una célula tardaría varios minutos en responder correctamente a cualquier cambio en la entrada! E incluso en ese caso la respuesta sería errónea, porque los pequeños errores de la suma se irían acumulando a lo largo de la cadena. Sin embargo en el ordenador esta técnica funciona muy bien, porque para cada paso sólo se

necesita aproximadamente un microsegundo y todos los cálculos se llevan a cabo sin errores.

Nuestros programas de visión para robots están llenos de atajos de este tipo, con los que se aprovechan la gran velocidad y la precisión con que calcula el ordenador. Por otro lado, los sistemas nerviosos están plagados de interconexiones que se solapan y que aprovechan el poder que tiene la maquinaria de construcción genética para reproducirse a sí misma.

Quizá todas estas triquiñuelas funcionen para la retina, pero no existe ninguna garantía de que también lo hagan para las distintas estructuras del cerebro. Es posible que algunas partes del cerebro utilicen sus neuronas de una forma tan sumamente inteligente que un programa de ordenador no pueda hacer otra cosa mejor que simular neuronas y sinapsis. Pero no es muy probable que sea así. El ejemplo de la retina sirve para ilustrar dos principios generales. El primero de ellos es que la reducida velocidad de conmutación y la limitada exactitud de las señales excluyen ciertas soluciones para los circuitos neurológicos que son fáciles para los ordenadores. El segundo es que si se aplica repetidamente una función suave a entradas que se solapan, la función se puede descomponer en partes de forma que esas subpartes se puedan usar más de una vez. Muchas de las estructuras neurológicas del cerebro tienen conexiones transversales de muchas entradas con muchas salidas, lo que las hace candidatas a ese tipo de economía. La corteza del cerebro humano, una de las estructuras mayores, es un disco arrugado de unos 2 mm de espesor y 20 cm de diámetro, que contiene diez mil millones de neuronas distribuidas en media docena de capas y «enhebradas» de forma bastante repetitiva. De esta lámina se ha estudiado bien el uno o dos por ciento, que rige la visión y se encarga de llevar adelante el proceso comenzado en las retinas y utilizando, por lo que parece, métodos semejantes. En las primeras capas se identifican bien los bordes y el movimiento en distintas direcciones y las capas de alimentación que responden a modelos más complejos, tales como las esquinas.

En el sistema nervioso se espera que exista algún tipo de simetría, porque la información contenida en los 1010 *bits* del conjunto de cromosomas haploides de un gameto o núcleo es insuficiente para «enhebrar» muchas de las 1014 sinapsis del cerebro. Lo que resulta interesante es constatar que este argumento no es válido para la Aplysia, tan estudiada, que tiene unas 100.000 neuronas agrupadas en 100 ganglios. Se ha podido hacer un diagrama de algunos de estos ganglios y parece que las neuronas y sus interconexiones son exactamente las mismas en todos los animales, y que cada una de las uniones juega un papel único y fundamental en el comportamiento del animal. Es verosímil que los pocos miles de millones de *bits* del código genético de la Aplysia contengan instrucciones especiales para «enhebrar» cada una de sus varios millones de sinapsis. Si resultara que era necesaria una simulación directa para esas partes del cerebro de los vertebrados que son especialmente irreductibles, entonces sería posible seguir mi trayectoria del tiempo. Si se obliga a un ordenador multiuso a que simule a las neuronas individuales, se observa que está mil veces por debajo de mi cifra de conversión para la retina. Sin embargo, se puede volver a aumentar la velocidad a costa de la flexibilidad, construyendo máquinas especiales para la simulación de neuronas en las que se utilice la misma cantidad de circuitos que el ordenador multiuso. Pero apuesto a que no será necesario.

Existe una asimetría cuando se compara la retina con un ordenador. ¿No es cierto que un ordenador programado para que imite un circuito neurológico tiene un potencial considerable que no se puede encontrar en el sistema neurológico? Porque, después de todo, el ordenador es multiuso y se lo puede reprogramar para que realice funciones totalmente distintas. Y la retina está condenada a hacer siempre los mismos cálculos. La diferencia sólo radica en la comodidad y en la velocidad de reprogramación. La retina *se puede* reprogramar, como ha sucedido muchas veces en el curso de nuestra evolu-

ción. El cálculo que lleva a cabo la retina de un cierto organismo concreto está previamente fijado, de la misma manera que los cálculos que lleva a cabo un ordenador los dicta el programa que contiene en ese momento. El grupo de todos los programas con los que puede funcionar un ordenador es semejante al grupo de maneras posibles en que se puede conectar una cantidad dada de tejido neurológico. La evolución seleccionó una cierta configuración de las propiedades e interconexiones de las neuronas, de la misma manera que nuestro estudio selecciona un cierto programa. Es evidente que existe una diferencia en el tiempo de programación. La configuración neurológica la controlan las instrucciones genéticas y cada cambio exige el desarrollo de un nuevo organismo, proceso que puede llevar años. En el ordenador, el cambio correspondiente se puede producir en unos minutos, es decir, un millón de veces más rápido.

Esto representa una gran ventaja, y creo que es una de las razones fundamentales por las que los ordenadores de uso especial llegarán a ocupar un lugar prominente en el campo de la robótica sólo después de que se haya terminado la investigación básica. Un ordenador de uso especial es un conjunto de aritmética, memoria y circuitos de control configurado de forma óptima para que desempeñe una cierta función. Puede ser mil veces más rápido que un ordenador multiuso, de tamaño y precio parecidos, que lleve a cabo la misma tarea. Pero el tiempo que se necesita para diseñar y construir una máquina tan especializada es aproximadamente el mismo que el que precisa un nuevo organismo para desarrollarse. En consecuencia, creo que la investigación se llevará a cabo en máquinas multiuso pero, una vez que se hayan entendido bien los requisitos de la equivalencia humana, se podrán construir máquinas de pensamiento especializadas que serán mucho más económicas que las que anticipo. Por otro lado, también creo que las posibilidades intrínsecas que tiene una máquina para perfeccionarse a sí misma son demasiado valiosas como para dejarlas de lado. En un robot maduro e inteligente, lo más probable es que haya algún mecanismo especial para sostener la superestructura multiuso.

¿No son las neuronas unos mecanismos altamente complejos y perfeccionados, producto de mil millones de años de evolución? ¿No es muy improbable que las podamos mejorar? No. En primer lugar, una parte importante del mecanismo de las neuronas está relacionado con su propio crecimiento y construcción de dentro afuera. Los componentes de los ordenadores actuales y futuros prescinden de esto, ya que están construidos desde el exterior. Lo cual supone una gran ventaja, ya que se puede utilizar toda la estructura para controlar la acción, la percepción y el pensamiento. En segundo lugar, parece que el mecanismo básico de intercambio de información de las neuronas —la emisión de productos químicos que afectan a las membranas exteriores de otras células— es tan primitivo que se puede observar incluso en las bacterias más simples que nadan libremente. Parece que los animales tienen que limitarse a esta disposición debido a las limitaciones en su proceso de diseño. La evolución darwiniana perfecciona inexorablemente un diseño dado, empuja ligeramente los parámetros por aquí y por allá, añade un paso aquí y elimina otro allí de forma lenta y chapucera. Sin embargo, no se puede hablar con propiedad de un nuevo diseño. Está fuera de su alcance hacer cambios fundamentales básicos porque habría que modificar correctamente muchas cosas a la vez. En contraste, los diseñadores humanos son muy competentes por lo que se refiere a mantener la forma general de una idea, aunque cambien todas sus partes. En su momento, las calculadoras se hicieron con ruedas dentadas y levas, después de relés, posteriormente de válvulas de vacío y, por último, de transistores y circuitos integrados. Es posible que pronto fluyan por sus cables luz o supercorrientes. En todo ese tiempo, las operaciones fundamentales han sido básicamente las mismas y los principios de diseño y el software que se ha creado para un cierto tipo de hardware se pueden pasar al siguiente con toda facilidad.

En la figura ***Potencia informática y memoria comparativa*** *del* ***Capítulo II****, ¿cuáles eran las presunciones relativas a la*

posición del sistema nervioso de los animales? Los datos son los que aparecen en la Tabla 1.

Tabla 1. *Sistemas nerviosos*

Animal	***Masa cerebral gramos***	***Neuronas***	***Potencia bits/seg***	***Capacidad bits***
Caracol		10^5	10^8	10^8
Abeja		10^6	10^9	10^9
Colibrí	0,1	10^7	10^{10}	10^{10}
Ratón	1	10^8	10^{11}	10^{11}
Visión humana	100	10^{10}	10^{13}	10^{13}
Ser humano	1.500	10^{11}	10^{14}	10^{14}
Elefante	3.000	2×10^{11}	2×10^{14}	2×10^{14}
Cachalote	5.000	5×10^{11}	5×10^{14}	5×10^{14}

A2

CÓMO SE MIDE LA POTENCIA DE UN ORDENADOR

La comparación entre la potencia relativa y el costo/efectividad de los distintos ordenadores ha sido siempre objeto de controversia, debido al efecto que tiene sobre las ventas. Pero el factor de desacuerdo que existe entre los fabricantes sobre la potencia de las máquinas de los demás es, por lo general, inferior a diez. Esta cifra tan pequeña no tiene ninguna repercusión ni sobre los diagramas ni sobre las conclusiones del Capítulo II, en el que las escalas son de billones. Y, sin embargo, cualquier fórmula particular para calcular la potencia puede llevar a error si se aplica a un contraejemplo desafortunado o diabólico. Por ejemplo, si la potencia de un ordenador se definiera, simplemente, como el número de sumas por segundo que es capaz de hacer, entonces un circuito especial compuesto por una serie de dispositivos para sumar, cuyo precio sería de pocos cientos de dólares y que no serviría para otra cosa, estaría por encima de un superordenador de diez millones de dólares. Mi intuición me ha sugerido otra manera de realizar esta medida, que a lo mejor es más complicada pero también más segura. Pueden tener la certeza de que, para máquinas normales, mis fórmulas proporcionan casi los mismos resultados que otros métodos más simples.

Las cosas con las que se hacen cálculos a gran escala pueden modificar sus variables internas y sus salidas de formas inesperadas. Podemos afirmar que una piedra inmóvil o que ruede, o incluso el ejemplo que se describe en el párrafo anterior, no precisan de muchos cálculos porque son bastante predecibles, mientras que un ratón que corretea por un laberinto necesita algunos más. La teoría de la información de Claude Shannon se basa en la manera de cuantificar la cantidad de *sorpresa* o de información que contiene el mensaje. Cuanto más

inesperada sea la parte siguiente del mensaje, más cantidad de información contiene. Para medir la información que proporciona un cálculo, utilizaré este método. La potencia de cálculo se define como la cantidad de información, o de sorpresa, que aparece por segundo cuando la máquina está funcionando, o sea, sus cambios continuos de un estado interno a otro. Cuanto más inesperado es el siguiente estado interno de la máquina, mayor es la cantidad de información que proporciona la transición a ese estado. Cuantitativamente hablando, si por lo que sabemos existe una probabilidad *p* de que la máquina pase a un cierto estado, entonces si lo hace ha comunicado $-\log_2 p$ *bits* de información ($\log_2$ significa logaritmo en base 2, y los *bits* son dígitos binarios). Si la probabilidad *p* fuera de 1/2, la transición comunicaría sólo 1 *bit* de información. En el caso de que la probabilidad fuera de $1/2^n$, entonces comunicaría n *bits*. Una *p* de 1/1.000 proporciona unos 10 *bits* de información.

La información media que comunica una transición se calcula multiplicando la información que se comunica en una transición a cada uno de los posibles estados siguientes por la probabilidad de pasar a ese estado y, después, se hace la suma de todas las posibilidades que existen. Es decir:

$$Información\ por\ transición = \sum_{i=1}^{N} -p_i \log_2 p_i\ bits$$

donde *N* es el número de estados posibles y p_i es la probabilidad de que el siguiente estado sea el *i*-ésimo. El valor calculado de una transición dada puede ser distinto para distintos observadores, ya que éstos asignarán probabilidades diferentes a los resultados. La información alcanza un máximo de $\log_2 N$ para un observador totalmente ignorante, que considere que todos los p_i son iguales. En el otro extremo, un testigo instruido podrá tener la certeza de que el estado siguiente será;, con lo que p_i=1 y todos los otros $p_i = 0$, siendo la información 0. O sea, que los cálculos que hace la máquina sólo le resultarán útiles si usted no conoce todas las respuestas de antemano.

Los cálculos precisan largas secuencias de este tipo de transiciones de un estado a otro. La *capacidad de información* total de un sistema es $\log_2$ de todos los estados a los que puede pasar. En un ordenador multiuso, este valor es simplemente el tamaño total de la memoria. Una máquina más potente puede pasar rápidamente por todos los estados. La *potencia de proceso* se mide dividiendo la información de la transición por el tiempo medio que requiere una transición. Esto nos da la siguiente fórmula:

$$Potencia = \frac{\sum -p_i \log_2 p_i}{\sum p_i t_i}$$

Las unidades son *bits* por segundo. Esta medida es reducida también por el grado de predictibilidad.

Las fórmulas anteriores encierran un cierto número de ideas. Un ordenador que repite una y otra vez un programa es totalmente predecible, por lo que su potencia de cálculo es cero. Los programas escritos en lenguajes de alto nivel o que precisan de medios interactivos con frecuencia funcionan con mucha más lentitud que si hubieran estado escritos directamente en el lenguaje de máquina del ordenador. Las estructuras de los lenguajes de alto nivel se traducen a las secuencias estereotipadas de las instrucciones de la máquina. De esta manera, el programa es más predecible que si estuviera escrito en puro lenguaje de máquina, con lo que se reduce la potencia de proceso efectiva. Si se amplía la memoria del ordenador, aumenta ligeramente su potencia sin que se produzca al tiempo un incremento en su velocidad no modificada. Entre las técnicas que existen para utilizar la memoria como forma de aumentar los cálculos se cuentan las tablas de resultados anteriores y las estructuras de reorganización de datos, que ocupan mucho espacio pero tienen la ventaja de que el acceso es más rápido. Éste efecto de que la memoria incremente la potencia de cálculo aparece en mi fórmula porque, en una instrucción del ordenador, la identidad del punto de la memoria al que se hace referencia supone tanta sorpresa como la operación que se va a rea-

lizar. A medida que aumenta el número de puntos posibles, aumenta la cantidad de sorpresa, aunque sea ligeramente. Doblar la memoria implica que la potencia aumente en un *bit* por tiempo de instrucción.

En las máquinas que están altamente en paralelo, especialmente en las que cuentan con una corriente de instrucciones simple para controlar numerosas unidades de proceso, la mayor parte de la sorpresa reside en los datos en paralelo, no en las instrucciones. Aunque el número total de *bits* que fluye por las corrientes de datos represente un límite superior para la potencia de proceso de este tipo de máquina, la potencia real puede ser considerablemente inferior debido a estados intrínsecos redundantes o predecibles. Es difícil calcularlos en máquinas diferentes diseñadas para realizar análisis de elementos finitos, procesos simbólicos, autómatas celulares o para jugar al ajedrez. Pero esta dificultad no afecta mucho a el gráfico del tema **Comparar ordenadores** del **Capítulo II**, ya que casi todas las máquinas que aparecen en ella son del tipo convencional Von Neumann, que sólo procesan un dato por instrucción. En las versiones futuras de este gráfico habrá que tener en cuenta las máquinas que se encuentran en paralelo masivamente, algunas de las cuales empiezan ahora a enfrentarse con problemas reales.

Incluso en el caso de las estructuras convencionales, las medidas de la potencia se confunden cuando se aplican las fórmulas a los ordenadores reales. ¿Cómo se pueden asignar las probabilidades a los distintos tipos de instrucciones si cada programa tiene su propio sistema de estadística? En los ordenadores con grandes grupos de instrucciones, muchas de las operaciones no se usan casi nunca. Además, para muchas máquinas antiguas, es difícil conseguir instrucciones detalladas. Yo hago la suposición de que cada máquina utiliza treinta y dos operaciones distintas (equivalentes a seis *bits)* que están mezcladas en la misma proporción. Si todos los puntos de la memoria tienen las mismas probabilidades de que una instrucción se dirija a ellos, entonces la información con la que contribuyen es igual

al logaritmo del tamaño de la memoria. Éste será el límite superior. Como puede cambiar el contenido de los propios puntos de la memoria, los datos almacenados son también una fuente de sorpresa, pero sólo en el caso de que se lean en vez de borrarlos y almacenar datos nuevos. Si suponemos que la mitad de las instrucciones leen datos, este canal contribuye con una información máxima de la mitad del tamaño de una palabra. En una máquina en paralelo controlada por una corriente simple de instrucciones, se debe tener en cuenta el tamaño global de la palabra de las corrientes de datos paralelos, que será el componente principal de la información total.

Otro de los puntos fundamentales es el tiempo. Y, una vez más, es difícil poder obtener muchos detalles. Para hacer una suma y una multiplicación se necesitan dos números. La suma es una de las operaciones típicas más rápidas del ordenador, mientras que la multiplicación es lenta. Parto del supuesto de que la instrucción global contiene siete instrucciones que duran el tiempo de una suma, y una que dura el tiempo de una multiplicación.

Con estas aproximaciones, la fórmula de la potencia se convierte en:

$$Potencia = \frac{6 + \log_2 memoria + palabra/2}{(7 \times T_{suma} + T_{multiplicar})/8}$$

donde *memoria* es la capacidad de la memoria rápida de la máquina en palabras a las que nos podemos dirigir individualmente y *palabra* es el tamaño de una palabra de datos en *bits*. La *capacidad* de la máquina se calcula multiplicando *memoria* por *palabra*. Para máquinas decimales, el número de *bits* se puede aproximar multiplicando por cuatro el número de cifras decimales. Esta fórmula es la que se ha utilizado para calcular los puntos de la figura del tema **Comparar ordenadores** del **Capítulo II** y los datos son los que aparecen en la Tabla 2.

UNA METÁFORA NÁUTICA

He definido dos cualidades esenciales para realizar cálculos interesantes o pensamientos. Son la *potencia* y la *capacidad*. Básicamente, la potencia es la velocidad de la máquina, y la capacidad es el tamaño de la memoria. El proceso de calcular se puede comparar con un viaje por mar en lancha motora. La duración del viaje dependerá de la potencia del motor de la lancha. El recorrido máximo queda limitado por la capacidad del tanque de gasolina. La velocidad efectiva disminuye, en general, si la trayectoria de la motora está fijada de antemano, por ejemplo si debe tomar la dirección este/oeste o norte/sur hacia su punto de destino en vez de ir en línea recta. Algunos cálculos son como un viaje a una ciudad conocida situada en una costa lejana. Otros, por el contrario, son como buscar sin mapa una isla perdida. Hacer cálculos en paralelos es como poseer una flota de botes: está bien para llevar a cabo búsquedas o para llegar a varios sitios, pero no sirve de gran cosa si de lo que se trata es de alcanzar rápidamente una meta alejada. Las máquinas de uso específico llevan un motor más grande para que el control sobre el timón sea menor. Añadirle a un ordenador discos y cintas es como ponerle a la motora más tanques de gasolina. La capacidad y, por lo tanto, la distancia que puede recorrer, aumentan, pero si los tubos de conexión son demasiado delgados, la velocidad del flujo de gasolina quedará limitada y, en consecuencia, la potencia efectiva del motor. Los dispositivos de entrada y salida son como las velas. Atrapan la potencia y la capacidad que existen en el entorno. De acuerdo con nuestra definición, la información externa es una fuente de variabilidad y, por lo tanto, de potencia. En términos más concretos, puede contener respuestas que, en otro caso, habría que calcular. Asimismo, el medio externo puede funcionar como memoria añadida, con lo que se incrementa la capacidad.

Tabla 2. *Máquinas calculadoras por año*

Coste *$ 1988*	Memoria *palabras*	Palabra *bits*	T_{sum} *seg*	T_{mult} *seg*	Potencia *bits/seg*	Capacid. *bits*	Pot/Coste *b/s/$*
Humano							
1×10^{5}	2×10^{1}	40	6×10^{2}	6×10^{2}	5×10^{-2}	8×10^{2}	5×10^{-7}
1891 — Ohdner (mecánica)							
1×10^{5}	6×10^{-2}	20	1×10^{2}	6×10^{-2}	7×10^{-2}	1×10^{0}	5×10^{-7}
1900 — Steiger Millonario (mecánica)							
1×10^{5}	1×10^{-1}	24	5×10^{1}	1×10^{2}	3×10^{-1}	3×10^{0}	2×10^{-6}
1908 — Tabuladora Hollerith (mecánica)							
5×10^{5}	8×10^{1}	30	5×10^{1}	2×10^{2}	4×10^{-1}	2×10^{3}	7×10^{-7}
1910 — Motor analítico (mecánica)							
9×10^{6}	1×10^{3}	200	9×10^{0}	6×10^{1}	8×10^{0}	2×10^{5}	8×10^{-7}

Coste *$ 1988*	Memoria *palabras*	Palabra *bits*	T_{sum} *seg*	T_{mult} *seg*	Potencia *bits/seg*	Capacid. *bits*	Pot/Coste *b/s/$*
1911 — Calculadora Monroe (mecánica)							
4×10^{5}	1×10^{0}	24	3×10^{1}	1×10^{2}	4×10^{-1}	2×10^{1}	1×10^{-6}
1919 — Tabuladora IBM (mecánica)							
1×10^{5}	5×10^{0}	40	5×10^{0}	2×10^{2}	8×10^{-1}	2×10^{2}	9×10^{-6}
1920 — Aritmómetro Torres (relés)							
1×10^{5}	2×10^{0}	20	1×10^{1}	1×10^{2}	7×10^{-1}	4×10^{1}	7×10^{-6}
1928 — National-Ellis 3.000 (mecánica)							
1×10^{5}	1×10^{0}	36	1×10^{1}	6×10^{1}	1×10^{0}	4×10^{1}	1×10^{-5}
1929 — Burroughs Clase 16 (mecánica)							
1×10^{5}	1×10^{0}	36	1×10^{1}	6×10^{1}	1×10^{0}	4×10^{1}	1×10^{-5}
1938 — Zuse-1 (mecánica)							
9×10^{4}	2×10^{1}	16	1×10^{1}	1×10^{2}	8×10^{-1}	3×10^{2}	1×10^{-5}
1939 — Zuse-2 (relés y mecánica)							
9×10^{4}	1×10^{1}	16	1×10^{0}	1×10^{1}	8×10^{0}	3×10^{2}	1×10^{-4}
1939 — BTL Modelo 1 (relés)							
4×10^{5}	4×10^{0}	8	3×10^{-1}	3×10^{-1}	4×10^{1}	3×10^{1}	9×10^{-5}
1941 — Zuse-3 (relé y mecánica)							
4×10^{5}	6×10^{1}	32	5×10^{-1}	2×10^{0}	4×10^{1}	2×10^{3}	1×10^{-4}
1943 — BTL Modelo 2 (relés)							
3×10^{5}	5×10^{0}	20	3×10^{-1}	5×10^{0}	2×10^{1}	1×10^{2}	6×10^{-5}
1943 — Colossus (válvulas)							
6×10^{5}	2×10^{0}	10	2×10^{-4}	2×10^{-2}	4×10^{3}	2×10^{1}	7×10^{-3}
1943 — BTL Moldelo 3 (relés)							
1×10^{6}	2×10^{1}	24	3×10^{-1}	1×10^{0}	6×10^{1}	4×10^{2}	4×10^{-5}
1944 — ASCC (Mark 1) (relés)							
2×10^{6}	7×10^{1}	70	3×10^{-1}	6×10^{0}	5×10^{1}	5×10^{3}	2×10^{-5}
1945 — Zuse-4 (relés)							
3×10^{5}	6×10^{1}	32	5×10^{-1}	2×10^{0}	4×10^{1}	2×10^{3}	1×10^{-4}
1946 — BTL Modelo 5 (relés)							
3×10^{6}	4×10^{1}	28	3×10^{-1}	1×10^{0}	7×10^{1}	1×10^{3}	2×10^{-5}
1946 — ENIAC (válvulas)							
3×10^{6}	2×10^{1}	40	2×10^{-4}	3×10^{-3}	6×10^{4}	8×10^{2}	2×10^{-2}
1947 — Harvard Mark 2 (relés)							
1×10^{6}	1×10^{2}	40	2×10^{-1}	7×10^{-1}	1×10^{2}	4×10^{3}	9×10^{-5}

Coste *$ 1988*	Memoria *palabras*	Palabra *bits*	T_{sum} *seg*	T_{mult} *seg*	Potencia *bits/seg*	Capacid. *bits*	Pot/Coste *b/s/$*
1948 — IBM SSEC (válvulas y relés)							
2×10^{6}	8×10^{0}	48	3×10^{-4}	2×10^{-2}	1×10^{4}	4×10^{2}	6×10^{-3}
1949 — EDSAC (válvulas)							
4×10^{5}	5×10^{2}	35	3×10^{-4}	3×10^{-3}	5×10^{4}	2×10^{4}	1×10^{-1}
1950 — SEAC (válvulas)							
3×10^{6}	1×10^{3}	45	2×10^{-4}	2×10^{-3}	8×10^{4}	5×10^{4}	2×10^{-2}
1951 — UNIVAC I (válvulas)							
4×10^{6}	1×10^{3}	44	1×10^{-4}	2×10^{-3}	1×10^{5}	4×10^{4}	3×10^{-2}
1952 — Zuse-5 (relés)							
4×10^{5}	6×10^{1}	32	1×10^{-1}	5×10^{-1}	2×10^{2}	2×10^{3}	5×10^{-4}
1952 — IBM CPC (válvulas y relés)							
4×10^{5}	9×10^{0}	144	8×10^{-4}	1×10^{-2}	4×10^{4}	1×10^{3}	9×10^{2}
1953 — IBM 650 (válvulas)							
8×10^{5}	1×10^{3}	40	7×10^{-4}	1×10^{-2}	2×10^{4}	4×10^{4}	2×10^{-2}
1954 — EDVAC (válvulas)							
2×10^{6}	1×10^{3}	44	9×10^{-4}	3×10^{-3}	3×10^{4}	5×10^{4}	2×10^{-2}
1955 — Whirlwind (válvulas)							
8×10^{5}	2×10^{3}	16	2×10^{-5}	3×10^{-5}	1×10^{6}	3×10^{4}	2×10^{0}
1955 — Librascope LGP-30 (válvulas)							
1×10^{5}	4×10^{3}	30	3×10^{-4}	2×10^{-2}	1×10^{4}	1×10^{5}	1×10^{-1}
1955 — IBM 704 (válvulas)							
8×10^{6}	8×10^{3}	36	1×10^{-5}	2×10^{-4}	1×10^{6}	3×10^{5}	1×10^{-1}
1959 — IBM 7090 (transistores)							
1×10^{7}	3×10^{4}	36	4×10^{-6}	2×10^{-5}	7×10^{6}	1×10^{6}	6×10^{-1}
1960 — IBM 1620 (transistores)							
7×10^{5}	2×10^{4}	5	6×10^{-4}	5×10^{-3}	2×10^{4}	1×10^{5}	3×10^{-2}
1960 — DEC PDP-1 (transistores)							
5×10^{5}	8×10^{3}	18	1×10^{-5}	2×10^{-5}	2×10^{6}	1×10^{5}	5×10^{0}
1961 — Atlas (transistores)							
2×10^{7}	4×10^{3}	48	1×10^{-6}	5×10^{-6}	3×10^{7}	2×10^{5}	2×10^{0}
1962 — Burroughs 5000 (transistores)							
4×10^{6}	2×10^{4}	13	1×10^{-5}	4×10^{-5}	2×10^{6}	2×10^{5}	5×10^{-1}
1964 — DEC PDP-6 (transistores)							
1×10^{6}	2×10^{4}	36	1×10^{-5}	2×10^{-5}	3×10^{6}	6×10^{5}	3×10^{0}

Coste *$ 1988*	**Memoria** *palabras*	**Palabra** *bits*	T_{sum} *seg*	T_{mult} *seg*	**Potencia** *bits/seg*	**Capacid.** *bits*	**Pot/Coste** *b/s/$*
1964 — CDC 6600 (transistores)							
2×10^{7}	5×10^{5}	64	3×10^{-7}	5×10^{-7}	2×10^{8}	3×10^{7}	1×10^{1}
1965 — IBM 1130 (chip híbrido)							
4×10^{5}	8×10^{3}	16	8×10^{-6}	4×10^{-5}	2×10^{6}	1×10^{5}	6×10^{0}
1966 — IBM 360/75 (chip híbrido)							
2×10^{7}	2×10^{6}	32	8×10^{-7}	2×10^{-6}	5×10^{7}	6×10^{7}	3×10^{0}
1967 — IBM 360/65 (chip híbrido)							
1×10^{7}	1×10^{6}	32	2×10^{-6}	3×10^{-6}	2×10^{7}	3×10^{7}	2×10^{0}
1968 — DEC PDP-10 (circuito integrado)							
2×10^{6}	1×10^{5}	36	2×10^{-6}	1×10^{-5}	1×10^{7}	5×10^{6}	8×10^{0}
1969 — CDC 7600 (transistores)							
3×10^{7}	1×10^{6}	64	1×10^{-7}	2×10^{-7}	5×10^{8}	6×10^{7}	2×10^{1}
1970 — GE-635 (transistores)							
6×10^{6}	1×10^{5}	32	2×10^{-6}	1×10^{-5}	1×10^{7}	4×10^{6}	2×10^{0}
1971 — SDS 920 (transistores)							
3×10^{5}	6×10^{4}	32	2×10^{-5}	3×10^{-5}	2×10^{6}	2×10^{6}	7×10^{0}
1972 — IBM 360/195 (chip híbrido)							
2×10^{7}	1×10^{5}	32	1×10^{-7}	2×10^{-7}	3×10^{8}	4×10^{6}	1×10^{1}
1973 — Data General Nova (circuito integrado)							
3×10^{4}	8×10^{3}	16	2×10^{-5}	4×10^{-5}	1×10^{6}	1×10^{5}	5×10^{1}
1974 — IBM-370/168 (circuito integrado)							
4×10^{6}	3×10^{5}	32	2×10^{-7}	4×10^{-7}	2×10^{8}	8×10^{6}	4×10^{1}
1975 — DEC-KL-10 (circuito integrado)							
1×10^{6}	1×10^{6}	36	8×10^{-7}	2×10^{-6}	5×10^{7}	4×10^{7}	5×10^{1}
1976 — DEC PDP-11/70 (circuito integrado)							
3×10^{5}	6×10^{4}	16	3×10^{-6}	9×10^{-8}	8×10^{6}	1×10^{6}	3×10^{1}
1976 — Apple II (circuito integrado)							
6×10^{3}	8×10^{3}	8	1×10^{-5}	4×10^{-5}	2×10^{6}	6×10^{4}	3×10^{2}
1977 — Cray-1 (circuito integrado)							
2×10^{7}	4×10^{6}	64	2×10^{-8}	2×10^{-8}	3×10^{9}	3×10^{8}	2×10^{2}
1979 — DEC VAX 11/780 (microprocesador)							
3×10^{5}	2×10^{6}	32	2×10^{-6}	3×10^{-6}	2×10^{7}	6×10^{7}	8×10^{1}
1980 — Sun-1 (microprocesador)							
4×10^{4}	3×10^{5}	32	3×10^{-6}	1×10^{-5}	1×10^{7}	8×10^{6}	3×10^{2}

Coste *$ 1988*	**Memoria** *palabras*	**Palabra** *bits*	**T_{sum}** *seg*	**T_{mult}** *seg*	**Potencia** *bits/seg*	**Capacid.** *bits*	**Pot/Coste** *b/s/$*
1981 — CDC Cyber-205 (circuito integrado)							
1×10^7	4×10^6	,32	3×10^{-8}	3×10^{-8}	1×10^9	1×10^8	1×10^2
1982 — IBM PC (microprocesador)							
3×10^3	2×10^4	16	4×10^{-6}	2×10^{-5}	5×10^6	4×10^5	2×10^3
1982 — Sun-2 (microprocesador)							
2×10^4	5×10^5	32	2×10^{-6}	6×10^{-6}	1×10^7	2×10^7	6×10^2
1983 — Vax 11/750 (microprocesador)							
6×10^4	1×10^6	32	2×10^{-6}	1×10^{-5}	2×10^7	3×10^7	3×10^2
1984 — Apple Macintosh (microprocesador)							
2×10^3	3×10^4	32	3×10^{-6}	2×10^{-5}	8×10^6	1×10^6	3×10^3
1984 — Vax 11/785 (microprocesador)							
2×10^5	4×10^3	32	7×10^{-7}	1×10^{-6}	5×10^7	1×10^5	2×10^2
1985 — Cray-2 (circuito integrado)							
1×10^7	3×10^8	64	4×10^{-9}	4×10^{-9}	2×10^{10}	2×10^{10}	1×10^3
1986 — Sun-3 (microprocesador)							
1×10^4	1×10^6	32	9×10^{-7}	2×10^{-6}	4×10^7	3×10^7	4×10^3
1986 — DEC VAX 8650 (microprocesador)							
1×10^5	4×10^6	32	2×10^{-7}	6×10^{-7}	2×10^8	1×10^8	1×10^3
1987 — Apple Mac II (microprocesador)							
3×10^3	5×10^5	32	1×10^{-6}	2×10^{-6}	4×10^7	2×10^7	1×10^4
1987 — Sun-4 (microprocesador)							
1×10^4	4×10^6	32	2×10^{-7}	4×10^{-7}	2×10^8	1×10^8	2×10^4
1989 — Cray-3 (arseniuro de galio)							
1×10^7	1×10^7	64	6×10^{-10}	6×10^{-10}	1×10^{11}	6×10^8	1×10^4

A3

LOS LÍMITES EXTERNOS DE LA COMPUTACIÓN

¿MENTE SIN MÁQUINAS?

En el Capítulo IV propuse la idea de que la mente es un modelo que se puede impresionar en muchos tipos diferentes de cuerpos o de depósitos. Pero fui todavía más allá y afirmé que se podía representar una mente por medio de cualquiera de los modelos absolutamente diferentes que forman una clase infinita y que eran equivalentes solamente en un cierto sentido abstracto y matemático. Las experiencias subjetivas de una persona son una propiedad abstracta que comparten todos los modelos de esta clase, de tal manera que esa persona tendría las mismas sensaciones independientemente del modelo que se utilizara para representarla. Esto nos lleva a la cuestión de que si la mente es, al fin y al cabo, una abstracción matemática, ¿por qué razón precisa una forma física? ¿No es cierto que las propiedades matemáticas existen aunque no estén escritas en ninguna parte? ¿No existe el dígito de pi que hace el número mil millones incluso aunque no hayamos conseguido calcularlo? En este sentido, las relaciones matemáticas abstractas que son los sentimientos de una persona, ¿no existen incluso en ausencia del *hardware* adecuado para calcularlas? Esto es un antiguo acertijo filosófico y aunque creo que debe ser verdad, no veo cómo se puede sacar de él ninguna conclusión que tenga sentido, ya que lo que parece implicar es que todo lo que es posible, existe. Así que vamos a considerar una línea de pensamiento un poco menos profunda. Supongamos que hay un programa que describe a una persona y está escrito en un medio estático, como un libro. Un ser superinteligente, capaz de leer

y entender el programa, debería poder deducir el comportamiento futuro de la persona codificada en diversas situaciones posibles. La existencia en el pensamiento de un observador inteligente es fundamentalmente igual que en una simulación de ordenador, y ya hemos sugerido que una mente se puede codificar satisfactoriamente en un ordenador. Pero un ser inteligente puede hacer con la simulación otras cosas además de ponerla en práctica rígidamente. No necesita modelar con exactitud todos y cada uno de los detalles de lo que contempla, se puede saltar las partes aburridas y las conclusiones que le resulten evidentes, aproximar otros pasos y mezclar alternativas. Los autores de ficción humanos lo hacen todos los días cuando se inventan aventuras para los personajes de sus libros. Lo único que les diferencia de nuestro ser superinteligente es que la imaginación de este último funciona a un nivel de detalle tal que las personas que simula son totalmente reales. Lo mismo que un autor de ficción, este ser puede pensar invirtiendo el curso del tiempo. Puede seleccionar una conclusión y, luego, hacer el razonamiento inverso y decidir qué es lo que ha precedido a esa conclusión. Es posible que el ser superinteligente prefiera imaginarse ciertos tipos de situaciones y se las ingenie para conseguir que sus simulaciones mentales hagan que sucedan. O acaso pueda crear las suficientes extravagancias en el mundo simulado como para atraer la atención de la persona simulada, ya que no tiene que inventar detalles carentes de importancia para la simulación ni hacer que los acontecimientos produzcan conclusiones concretas.

Parece que la mecánica cuántica, una de las piedras sillares de la física moderna, implica que en el mundo real, tal y como lo conocemos, tienen lugar sucesos inobservados de todas las maneras posibles (lo cual equivale a decir que lo que ocurre no depende de ninguna decisión) y que la superposición de todas estas posibilidades tiene efectos observables, entre los que se cuentan las coincidencias misteriosas en épocas o lugares lejanos. ¿Existe alguna relación entre estas ideas? Una vez más, me he metido en un acertijo, aunque esta vez haya alguna posibilidad de encontrar, finalmente, una respuesta.

EL PENSAMIENTO NO DETERMINISTA

Los universos de los autómatas celulares que se esbozaron en el Capítulo VI tienen la rigidez de un mecanismo de relojería y nos resulta fácil imaginarlos. Pero la física moderna ha descubierto que nuestro propio universo tiene unos fundamentos mucho más interesantes. Simplemente pensar en lo que en realidad significa es un tema de especulación fascinante. Sólo se ha llegado al consenso en un punto: la verdad es muy *rara*. Lo que viene a continuación tiene algo de autocomplacencia, un intento de conducir esa «rareza» a un fin práctico. Pero, evidentemente, la realidad puede resultar ser todavía más rara.

Antes de que se inventaran los ordenadores, los cálculos científicos rigurosos eran difíciles y largos, y era fácil cometer errores. Los problemas de ingeniería práctica se abordaban de una forma específica que hacía que se perdiera generalidad, con tablas numéricas, métodos gráficos y dispositivos tales como, reglas de cálculo y, en este siglo, calculadoras mecánicas. Los ordenadores cambiaron por completo la situación. Aunque podían enfrentarse con problemas enormes, necesitaban que se les proporcionaran por anticipado indicaciones absolutamente precisas y detalladas, paso a paso, sobre el método a seguir. Elaborar esos programas y hacerlo bien era aburrido y lento. Sin embargo, era más lento hacer a mano todos los cálculos. Suponía una gran ventaja que se pudiera usar el mismo programa para varios problemas del mismo tipo y, con frecuencia, esa generalización servía para clarificar la estructura de los cálculos. Asimismo, daba lugar a que apareciera el nuevo campo matemático de los *cómputos complejos*, es decir, el estudio de las dificultades intrínsecas que implica computar las soluciones para distintos tipos de problemas.

Un programa general puede utilizar programas de diferentes tamaños y, por lo general, el tiempo de funcionamiento aumenta cuando crece el tamaño del problema. Por ejemplo, los programas más sencillos de distribución duran un tiempo proporcional al cuadrado del número de objetos a distribuir. Si se

dobla el número de objetos, el tiempo de distribución se multiplica por cuatro. Se han descubierto programas más complicados cuyos tiempos crecen con más lentitud, y los estudiantes de los cómputos complejos lograron demostrar que los programas de distribución más rápidos precisaban un número de pasos que era proporcional al logaritmo del número de objetos multiplicado por el número de objetos: para distribuir un millón de objetos se necesita aproximadamente diez mil veces el tiempo de distribución de mil objetos. El tiempo crece con más rapidez que el número de objetos, pero no llega a ser su cuadrado. Para otros tipos de problemas se demostró que la dificultad crecía con el cubo, con la cuarta potencia o con otra. Se dice que los problemas cuya solución es proporcional a alguna potencia fija del tamaño del problema son de tipo polinómico o tipo P. En la era de los ordenadores, se consideran fáciles.

Pero otras clases de cálculos, también importantes, resultaron ser mucho más difíciles. Un ejemplo es el denominado problema del *viajante*, que supone encontrar el camino más corto que pase por una serie de ciudades exactamente una vez. Las mejores soluciones exactas que se han hallado suponen que el programa tiene que verificar casi todas las rutas posibles y, luego, elige una. Las rutas posibles se enumeran calculando las permutaciones de las ciudades. Cada permutación se puede comprobar en un tiempo polinómico, pero el número de permutaciones crece exponencialmente y se multiplica cuando se añade una ciudad nueva. Para un número de ciudades relativamente pequeño, el problema es enormemente difícil. Un ordenador hipotético, capaz de explorar simultáneamente todos los caminos posibles (una simple abstracción matemática conocida con el nombre de máquina *no determinista*, porque no toma decisiones en las «bifurcaciones», sino que se divide en dos y sigue los dos caminos) sería capaz, en principio, de resolver el problema en un tiempo polinómico. Por esta razón, a esos problemas tan difíciles se les da el nombre de PN, por los polinomios no deterministas. Muchos problemas PN se pueden transformar matemáticamente en otros y, además, una solución rápida para uno de ellos (P) sirve para resolverlos todos. Pero,

por desgracia, no se ha encontrado ninguna solución rápida que funcione en un ordenador real y determinista, y no se sabe si existe esa solución. Esta cuestión, denominada ¿P = PN?, es fundamental porque algunos problemas de extremada importancia como, por ejemplo, el diseño del *hardware* y el *software* óptimos y el razonamiento automático, son PN. Como existen muchas soluciones exactas, y demasiado lentas, los sistemas de diseño actuales tienen que conformarse con los métodos aproximativos, que no garantizan que la respuesta final sea la mejor y que a veces proporcionan unos resultados bastante malos.



Pienso, luego existo
Un Descartes simulado deduce correctamente su propia existencia. Quién (o qué) lleve a cabo la simulación, no supone ninguna diferencia: el mundo simulado es completo en sí mismo.

Supongamos, ya que parece lo más probable, que no exista ningún atajo para resolver los problemas PN. No importa la velocidad a que funcione la máquina convencional que utilicemos para resolverlos; cualquier pequeño incremento en el tamaño del problema la dejará abrumada. Una inteligencia que realice diseños para su autoperfeccionamiento se encontrará con muchos problemas PN. La eficacia de sus diseños y, en consecuencia, su futuro, dependerán directamente de la forma en que se enfrente a estos problemas, así que los métodos heroicos quedan justificados.

REPRODUCCIÓN E INDETERMINISMO CUÁNTICO

Un ordenador no determinista, capaz de engendrar indefinidamente versiones de sí mismo para examinar respuestas alternativas, es una elaboración matemática. Se puede hacer una aproximación, hasta cierto punto, con un multiprocesador que contenga un número fijo de procesadores diferentes. Y, en el curso de los cálculos, cada vez que se llega a una bifurcación, se invoca un nuevo procesador. Con el tiempo, sin embargo, el último procesador estará ocupado y habrá que calcular las bifurcaciones posteriores de forma secuencial. La velocidad de los cálculos se puede incrementar tantas veces como procesadores individuales haya, y esto tiene poca influencia sobre la velocidad astronómica a que crecen los problemas PN. Pero ¿qué sucedería si se pudiera ir aumentando el número de procesadores a medida que creciera el tamaño del problema?

Para los miembros de cualquier raza que aspire a la inmortalidad, la capacidad de reproducción es esencial. Nuestras máquinas ya se reproducen, aunque los seres humanos biológicos sean imprescindibles en el proceso. A medida que el futuro esbozado en este libro se vaya haciendo realidad, las máquinas irán avanzando cada vez más hasta que se haga realidad la fabricación automática de máquinas. En el futuro superinteligente, la reproducción de pequeñas unidades de pensamiento será un juego de niños. Es, por lo tanto, razonable suponer que

los ordenadores pequeños podrán hacer cálculos y también copias de sí mismos. La reproducción, al fin y al cabo, queda limitada por la cantidad finita de materia y energía de que disponemos, pero los límites pueden ser astronómicamente altos si se producen máquinas muy pequeñas en un medio nutritivo energético muy amplio, por ejemplo, máquinas del tamaño de una bacteria en los océanos de Júpiter o en las nubes de polvo interestelares de las estrellas púlsar. Un ordenador así se reproduciría dando lugar a otros dos que, a su vez, darían lugar a cuatro, estos cuatro a ocho, los ocho a dieciséis y así sucesivamente. Este crecimiento exponencial daría lugar a un número astronómico que, en un cierto punto, sería igual al crecimiento exponencial de un problema PN. Si en el lapso de tiempo de una «generación» una máquina puede hacer una cierta cantidad de cálculos o bien reproducirse una vez, la mejor estrategia sería que se reprodujera como loca hasta que hubiera tantas máquinas como alternativas a examinar y que, entonces, cada una de ellas examinara una alternativa. Después de haber calculado las respuestas posibles, las máquinas se enzarzarían en una especie de torneo «la mía es mejor que la tuya». Dos máquinas compararían sus propias respuestas, elegirían el mejor resultado, lo adoptarían ambas y saldrían a competir con otras. La mejor respuesta se extendería rápidamente a todas las máquinas y todas conocerían el resultado final. En este proceso, la fase de reproducción dura un tiempo linealmente proporcional al tamaño del problema: la elaboración de las respuestas individuales tiene un tamaño polinómico y la respuesta de la fase del torneo, lo mismo que la reproducción, es de tamaño lineal. Esto permite resolver problemas PN de tamaño moderado en un tiempo polinómico. Pero los problemas muy grandes ocuparían todo él espacio disponible, con lo que quedan fuera de la cuestión. (La biosfera de la Tierra está llevando a cabo un proceso de este tipo.)

En realidad, los ordenadores no deterministas son una ficción matemática. ¿O no? La mecánica cuántica está impregnada de indeterminismo. En mecánica cuántica, las probabilidades de obtener un cierto resultado se predicen sumando todas

las formas indistinguibles en que puede producirse un suceso y elevando al cuadrado el resultado de esa suma. Esto tiene una consecuencia extraña: la existencia de otras posibilidades excluye ciertos resultados posibles. Un ejemplo excelente es el experimento de *las dos rendijas.* Una fuente puntual radia fotones de luz a una pantalla que tiene dos ranuras, como se puede ver en la figura de la página 218. Los que consiguen atravesar las rendijas se encuentran con un dispositivo de detección de fotones (por lo general, una película fotográfica, pero el ejemplo queda más claro si usamos sensores individuales que hacen un chasquido cuando la luz choca con ellos). Si la fuente de luz es tan débil que sólo emite un fotón cada vez, los sensores los detectan independientemente, unas veces unos y otras otros. Todos los fotones aterrizan exactamente en un lugar. Pero si se van contando los fotones que han aterrizado en cada detector, nos encontramos con un hecho inesperado. Algunos detectores no perciben ningún fotón, mientras que otros que están próximos perciben muchos, y los situados un poco más allá tampoco perciben ninguno. Si se cuenta el número de chasquidos de cada detector y se hace una tabla, el resultado es idéntico al del modelo de bandas de interferencia que habríamos obtenido si desde el punto donde está situada la fuente de luz se hubieran emitido ondas. Estas ondas habrían atravesado las dos rendijas y se habrían transformado en dos ondas que habrían interferido en la pantalla; la interferencia sería constructiva en algunos lugares y destructiva en otros, según se puede ver en la figura de esta misma página.

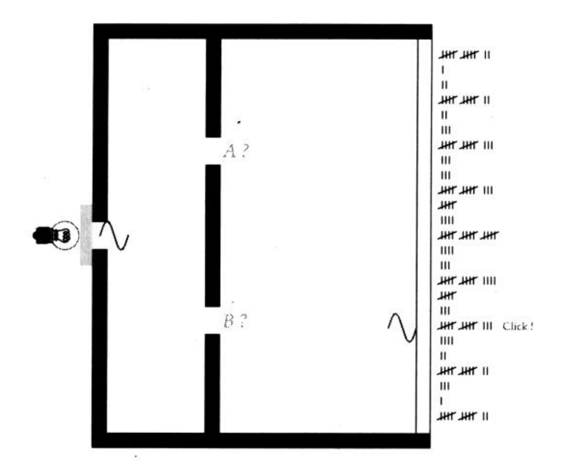
Pero ¿ondas de *qué*? Cada fotón sale de un lugar y aterriza en otro. ¿No está siempre en un cierto lugar en cada punto de su trayectoria? ¿No atraviesa una rendija o la otra? En ese caso, ¿cómo es posible que la presencia de otra rendija le *impida* aterrizar en un cierto lugar de la pantalla? Porque si se bloquea una de las rendijas, el número total de fotones que aterriza en la pantalla se divide por dos, pero ha desaparecido el fenómeno de interferencia, y algunos de los lugares en los que no caía ningún fotón con las dos rendijas abiertas empiezan a registrar choques. La respuesta de la mecánica cuántica es que, durante

el vuelo, la posición del fotón es desconocida y la define una onda compleja que describe todas las posiciones posibles. Esta onda fantasmal atraviesa *las dos* rendijas (aunque describe la posición de un solo fotón) e interfiere en la pantalla consigo misma, cancelándose en algunos puntos. Una vez allí, la onda toma una decisión y el fotón aparece sólo en uno de los puntos posibles. Esta condición de onda del fotón antes de que se produzca su choque con la pantalla recibe el nombre de *estado mixto* o *superposición de estados.* Y a la aparición repentina del fotón en un solo detector se la llama *colapso de la función de ondas.*

Esta explicación inquietaba muchísimo a algunos de los mismos físicos que habían contribuido a formular la teoría, en especial a Albert Einstein y a Erwin Schrödinger. Con el fin de formalizar sus objeciones intuitivas, idearon experimentos cuyos resultados contradecían la teoría. En algunos, una medida realizada en un cierto lugar producía el colapso inmediato de la función de ondas en un punto lejano: un efecto más rápido que la luz. En otro ejemplo, menos formal, al que se llamó el Gato de Schrödinger, una caída radiactiva que podía tener lugar (o no) en una caja sellada hacía (o impedía) que muriera un gato que estaba dentro de la caja. Schrödinger consideraba absurda la descripción que la teoría hacía de la caja cerrada: un estado mixto en el que el mismo gato estaba muerto y vivo. Lo que sugirió es que la teoría simplemente expresaba la ignorancia del observador, ya que, en el interior de la caja, no existía ninguna ambigüedad sobre la suerte del gato. Ésta es la teoría de las *variables ocultas,* es decir, el sistema tiene un estado definido en todo momento, pero algunas partes de él quedan temporalmente ocultas para los observadores.

Algunos de los resultados más «absurdos», aunque sean experimentales, los ha obtenido Alain Aspect, de la Universidad de París, a partir de experimentos muy ingeniosos y modernos. Estas demostraciones excluyen las teorías de variables ocultas más simples y naturales, las *locales,* en las que, por ejemplo, la

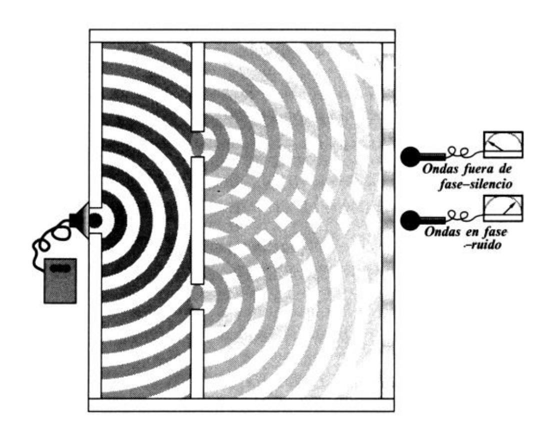
información oculta sobre cuál fue la rendija que atravesó el fotón está contenida en el propio fotón, o las que afirman que el estado de salud del gato de Schrödinger es parte del felino.



Experimento de las dos rendijas
Un fotón que escoja el detector en la pantalla S puede haber llegado hasta ella a través de la rendija A o de la B. No hay manera de distinguirlo. En mecánica cuántica se deben sumar las «amplitudes» de los dos casos. En algunos puntos de la pantalla, la suma es constructiva y es probable encontrar un fotón; en otros puntos cercanos, las amplitudes se cancelan y nunca se encontrarán fotones en ellos.

También encierran posibilidades las teorías de variables ocultas *no locales*, de acuerdo con las cuales la información no medida se distribuye por un espacio ampliado. Resulta muy sencillo construir teorías de este tipo, que dan los mismos resultados que la mecánica cuántica. Pero a muchos físicos no les interesan: ¿por qué introducir una explicación más complicada y con más variables cuando es suficiente con las ecuaciones normales y corrientes? Además, filosóficamente hablando, las teorías de variables ocultas son sólo un poco menos misteriosas que la propia mecánica cuántica. ¿Qué significa que la «posición exacta» de una partícula se encuentra distribuida por una

porción grande de espacio? Esta cuestión suscitó una viva controversia entre los fundadores de la mecánica cuántica a principios de siglo. Y recientemente ha vuelto a suscitar interés.



Dos rendijas y ondas
Cuando las dos rendijas son atravesadas por ondas sonoras, se produce el fenómeno de interferencia. Pero no se escuchan «clicks» individuales. Cada onda afecta suavemente a todos los detectores.

MUCHOS MUNDOS Y EL ORDENADOR PN. «DÍA DEL JUICIO FINAL»

En el experimento de las dos rendijas, un fotón que estaba destinado a llegar a una pantalla *tenía* que atravesar una rendija, la A o la B. El fenómeno de interferencias sugiere que, de alguna forma, se las arreglaba para atravesar las dos a la vez. En 1957, Hugh Everett, de Princeton, publicó en su tesis de doctorado una explicación basada en las variables ocultas no locales que, posiblemente, es la más extravagante de todo este rompecabezas. Según el modelo de Everett, el fotón atraviesa las dos rendijas *en universos diferentes*. En cada uno de los puntos en los que es necesario tomar una decisión, el universo, o por lo menos la porción más inmediata, se divide lo mismo

que las páginas que salen de una foto- copiadora. Hasta el momento en que se realiza la medida, las distintas versiones del universo permanecen muy próximas e interfieren unas con otras, produciéndose, por ejemplo, el fenómeno de bandas en una pantalla. Si una cierta medida hace que se pueda distinguir una posibilidad de otra, entonces el universo diverge. Y, a la inversa, la definición de la «medida» viene dada por la divergencia. La interferencia cesa y en cada universo, ahora independiente de los otros, una versión diferente del observador puede contemplar un resultado inequívoco y diferente.

Para expresar el número de universos distintos que, de acuerdo con esta idea, se crean a cada instante podríamos usar la palabra «astronómico», pero el calificativo se quedaría corto. Las cifras serían de tal magnitud que sólo imaginarlas resultaría aterrador, pero no nos debemos dejar intimidar por ellas. También se tuvieron que enfrentar con la sorpresa las primeras sugerencias sobre el número de átomos que había en una partícula de materia, la distancia a las estrellas más próximas o el tamaño y la edad del universo. Las objeciones de los físicos prácticos a la variedad de universos tienen otra base. Una vez que *se lleva a cabo* una medida, los universos en los cuales el resultado es distinto que en el nuestro ya no ejercen ninguna influencia sobre nosotros. Y postular que continúan existiendo supone una complicación innecesaria. Y, en ese caso, la idea de los muchos mundos se reduce a la mecánica cuántica convencional. Las variables ocultas (que identifican el universo *del que* usted está hablando) se eliminan porque son innecesarias.

El indeterminismo de la mecánica cuántica, ¿puede ser de alguna ayuda para resolver problemas difíciles? En otros campos se han empleado los denominados métodos holográficos, que utilizan luz láser coherente para buscar simultáneamente ciertas cosas (huellas dactilares o señales del radar enemigo, por ejemplo). Si se aplican al problema adecuado, los métodos holográficos son rápidos y eficaces y la acción se puede interpretar como el efecto de estados mezclados. También se pueden considerar de una forma más clásica, como la interferencia

de las ondas luminosas. Las ondas son muy interesantes, porque su propagación es como una especie de reproducción y se pueden usar distintas partes de un gráfico de ondas para realizar distintas partes de un cálculo. Pero como sólo se han propuesto o conseguido aceleraciones lineales, los problemas PN siguen siendo difíciles.

La idea de los muchos mundos tiene otra consecuencia. John Gribbin es escritor y físico y ha expresado las posibilidades más extravagantes en varias historias, artículos y libros. El argumento de una de sus historias es el siguiente.

EL MECANISMO DEL DÍA DEL JUICIO FINAL

Dos constructores de un futuro superacelerador de partículas (inmensamente costoso) tienen un problema. La máquina lleva meses terminada pero, hasta la fecha, todos los intentos para usarla han sido un fracaso. El problema no reside en el diseño sino, por lo que parece, en la mala suerte del diseñador. Un rayo provocó un apagón justo cuando la acababan de conectar, luego se fundió un fusible, un conserje tropezó con un cable, y un ligero temblor de tierra hizo que se pusieran en marcha los mecanismos de emergencia y se cortara la luz. Todos los incidentes fueron distintos y, aparentemente, no estaban relacionados.

Pero acaso todos ellos fueran un enorme golpe de suerte. Se hicieron nuevos cálculos y se descubrió que existía la posibilidad de que la máquina fuese tan potente que se podría producir un colapso del vacío de tal magnitud que éste pasaría a un estado energético inferior. Se podría producir una explosión cósmica que se propagaría a la velocidad de la luz desde el punto de colisión del acelerador y que, finalmente, destruiría todo el universo. ¡Por los pelos!

¿O no? Si se hubiera destruido el universo, no quedaría nadie para lamentarlo. Y ¿qué sucedería si la idea de los muchos mundos fuera correcta? La máquina habría funcionado en algunos universos. Para todas las cuestiones prácticas, esos

mundos habrían dejado de existir. Y en el que quedara, sólo habría un par de físicos perplejos rascándose la cabeza y preguntándose qué era lo que no había funcionado bien esta *vez. Habiendo tantos universos casi idénticos, no parece que la destrucción de unos pocos fuese a tener grandes consecuencias. Y se les ocurre una idea. ¿Por qué no reforzar los puntos débiles de la máquina de forma que sea altamente improbable que se produzca un fallo fortuito y luego conectarla a un detector de ataques nucleares, como la máquina del Día del Juicio Final que aparece en la película de Stanley Kubrick* Dr. Strangelove? *Si se produjera un ataque, se contestaría destruyendo el universo atacante. Sólo quedarían vivos los universos en los que, por alguna razón (el general sufrió un ataque al corazón, el sistema de lanzamiento de misiles no funcionó, el primer ministro tuvo un gesto de compasión...) no se hubiera producido el ataque. La máquina de la película* Strangelove *sólo era efectiva como arma disuasoria si el otro bando conocía su existencia. Pero no es así en la versión de los muchos mundos. No se puede producir ningún ataque (observable por alguien) mientras esté en funcionamiento, por muy secreta que sea su existencia.*

Evitar una guerra nuclear es, ciertamente, un objetivo laudable y justificaría la destrucción de un número incierto de posibles universos. Pero ¿podemos usar el mismo método para resolver problemas cotidianos? Volviendo a los problemas PN, ¿se podrían utilizar de alguna manera los universos paralelos para buscar posibles soluciones alternativas simultáneas que permitieran que un ordenador convencional funcionara como si fuera no determinista? Empecemos conectando a un ordenador un mecanismo para destruir el universo que se active con una cierta instrucción del ordenador. Mientras tanto, conectemos también un generador aleatorio de números basado en un siseo que provenga de un reos- tato caliente o de los chasquidos de un contador geiger. Las soluciones potenciales a los problemas PN vienen caracterizadas por un cierto número que deseamos minimizar. Por ejemplo, en el problema del viajante, la

longitud total del trayecto que tiene que recorrer. Suponga una longitud probable y escriba un programa para elegir un trayecto al azar (usando el generador que ha añadido). Haga que el programa active el dispositivo del día del juicio final si el trayecto generado al azar es mayor que el que usted había supuesto. Ponga en funcionamiento el programa. Lo más probable es que se encuentre con que el trayecto aleatorio que ha generado su programa es más corto que el que usted había supuesto, porque se han suprimido los universos en los que el trayecto era más largo. Pero supongamos que no *hay ningún trayecto* tan corto. Bueno, pues entonces su ordenador se debe haber desconectado y no ha destruido el universo.

Es un fastidio que el ordenador escape a su control y se le desconecte, así que, ¿por qué no coloca una conexión débil? Instale otro mecanismo, controlado térmica o radiactivamente, que no tenga muchas probabilidades de interrumpir los cálculos de su ordenador. Si activa el programa ahora, o bien le proporcionará una respuesta más corta que la que usted ha supuesto, o bien se interrumpirá antes de terminar. En el primer caso, la longitud del trayecto que habrá escogido será más corta. En el segundo, más larga. Vuelva a activar el programa. Repita el proceso hasta encontrar una longitud del trayecto L tal que la máquina pueda hallar respuesta cuando la suposición es L y se interrumpa cuando es L-l. La solución L es la respuesta óptima para su problema. Para que todo el proceso sea automático, la búsqueda de esta L se puede incorporar al propio programa. Existe una estrategia óptima para hacerlo. Recibe el nombre de *búsqueda binaria* y puede obtener el resultado en un número de pasos proporcional al logaritmo de L, es decir, simplemente en veinte pasos cuando L es de un millón. Cada uno de los pasos, que supone generar una respuesta aleatoria, necesita un tiempo polinómico. Por lo tanto, la respuesta a la pregunta es sí, una versión de los muchos mundos de la mecánica cuántica puede resolver en tiempos pequeños problemas PN arbitrariamente difíciles.

Sin embargo, el costo de los universos destruidos es pasmoso. Supongamos que queremos usar el mecanismo del día

del juicio final para evitar la guerra nuclear. Si la probabilidad anual de que se produjera un ataque nuclear fuera del 50%, cada año se destruirían la mitad de los universos posibles. Pero en el problema del viajante, si recorriera 100 ciudades, de 100^{100} (un 1 seguido de 200 ceros) soluciones sólo valdrían unas pocas. Por lo tanto, un ordenador que utilizara el mecanismo del día del juicio para resolver el problema, ¡destruiría aproximadamente 100^{100} mundos por cada uno que sobreviviera! La velocidad a que se producen los nuevos universos es mucho más elevada, así que a lo mejor no importa mucho.

En la actualidad, solamente hay un componente de esta solución del problema PN que no se puede construir inmediatamente: el activador. El diseño de Gribbin depende de un tipo de física altamente especulativa que ha perdido bastante popularidad. Pero ¿es realmente necesario destruir todo el universo si la respuesta que se obtiene no es la correcta? Evidentemente, no. Subjetivamente hablando, es tan necesario como destruirle a usted. Ahora bien, hoy en día se puede llegar a *ese* resultado. Por ejemplo, empalme la conexión del día del juicio de su ordenador a una carga explosiva que se colocará usted en la cabeza. Si utiliza el programa del viajante, lo más probable es que acabe volándose los sesos. Pero es rápido e indoloro. Y en uno de los 100^{100} mundos, usted tendrá la suerte de sobrevivir y de obtener la respuesta correcta. La carga explosiva que se había colocado en la cabeza estará intacta, lista para que usted aborde el próximo problema.

La idea anterior funcionaría para otros problemas de la vida. Si usted desea conseguir un resultado, aunque sea improbable, y fracasa, entonces usted se destruye a sí mismo. Pero sobrevivirá en algún universo y en él logrará su deseo. Si la interpretación de la mecánica cuántica de los muchos mundos es correcta, y es posible que lo sea, ¿por qué no es el suicidio una solución adecuada a los problemas cotidianos? Es posible que las demografías de los múltiples universos tengan alguna relación con la respuesta. Si a usted no le importa perder la vida, puede que haya algunas copias de usted por los universos.

Un universo elegido al azar estará poblado en general por individuos que luchan con éxito por evitar la muerte siempre que sea posible. Si los universos son realmente infinitos en número, esto apenas significa nada. Una billonésima de infinitud es la misma infinitud. Sin embargo, si el número de universos es simplemente enorme, las consecuencias pueden ser serias. Y si no se obtienen todos los resultados posibles, entonces lo más probable es que, cuando usted se destruya a sí mismo, acabe realmente con su existencia.

¿UN MUNDO Y NO MUCHOS?

Puede ser que la indeterminación de la mecánica cuántica sea simplemente una especie de ilusión y que sólo exista un mundo. A continuación le ofrecemos el esbozo de un modelo en el que las incertidumbres en cualquier punto, o las variables ocultas, son simplemente «ruido» del resto del universo.

Imagine que en alguna parte existe un volumen esférico uniformemente lleno de un gas compuesto por un número enorme, pero finito, de partículas en movimiento. Unas ondas de presión atraviesan el gas y se propagan a su velocidad del sonido, *s*. Suponga que no se puede enviar ninguna señal a mayor velocidad. La esfera tiene resonancias que corresponden a los trenes de ondas que atraviesan todo su volumen con distintos ángulos y frecuencias. Cada combinación de una dirección y una frecuencia concretas se llama *configuración de onda*. Existe una transformación matemática denominada *transformada de Fourier* (espacial) que se utiliza para ordenar las configuraciones de onda. La transformada de Fourier combina de distintas formas el fenómeno de las presiones que se encuentran en el volumen original de la esfera *(V)* y cuyo resultado es un nuevo grupo de valores esféricos *(F)*. En el centro de *F* se encuentra un número que representa la densidad media de *V*. En sus inmediaciones hay números (complejos) que proporcionan la intensidad de ondas, en distintas direcciones, cuyas longitudes de onda son exactamente el diámetro de *V*. Al doble de

la distancia del centro de *F* se encuentran las intensidades de configuraciones de onda que recorren dos ciclos en *V*. Están rodeadas por otra capa que contiene configuraciones cuya longitud de onda es un tercio del diámetro de *V*, y así sucesivamente. Cada punto de *F* describe una onda que llena *V* con una dirección y un número de ciclos dados por la orientación del punto y la distancia al centro de *F*. O, para expresarlo de otra manera: la dirección en *F* corresponde a la dirección en *V;* el radio de *F* es proporcional a la frecuencia en *V*. Como cada onda está compuesta por agrupaciones periódicas de partículas de gas, el espaciado entre las partículas fija el límite inferior de la longitud de onda, es decir, el límite superior de la frecuencia, y el límite del radio de la esfera *F*. Cuanto más próximas estén las partículas, más grande debe ser *F*.

Existe un teorema sobre la transformada de Fourier que afirma que si se incluyen frecuencias suficientemente altas, entonces *F* contiene aproximadamente tantos puntos como partículas hay en *V y* que, por lo tanto, toda la información necesaria para reconstruir *V* se encuentra en *F*. En realidad, *F y V* son simplemente distintas descripciones de la misma cosa y tienen la interesante propiedad de que cada partícula de *V* contribuye al valor de cada punto de *F, y* que cada uno de los puntos de *F* es un componente del movimiento de cada partícula de *V*.

Si las partículas de *V* chocan unas contra otras o interaccionan de forma no lineal, entonces la energía se puede transferir de una configuración de onda a otra, es decir, un punto de *F* puede ser más fuerte a costa de otro. Entre todas las configuraciones de onda se producirá una cierta transferencia aleatoria de energía. Además, habrá una interacción más sistemática entre configuraciones de onda «próximas»: las que tengan frecuencia y orientación parecidas y estén, por lo tanto, cercanas en el espacio *F*. Esas ondas, en *V*, estarán en fase gran parte de su longitud. Como el gas no es lineal, los agrupamientos periódicos de partículas de gas que producen una configuración influirán sobre la capacidad de agrupamiento de una configuración vecina que tenga un período semejante.

Resumiendo, se produce una interacción sistemática entre los puntos de *F* que están próximos y no en los que están alejados. Esta interacción se puede considerar como una física del mundo *F*. Si es suficientemente completa, puede sostener los cimientos de estructuras complejas, la vida y la inteligencia, como sucede con la nuestra. Imaginemos a un físico hecho de la materia de *F* para el cual los puntos de F son simplemente lugares, no complicadas funciones de otro espacio. Podemos deducir algunas de las «leyes de la física» que este ser descubriría razonando sobre los efectos en *V* y trasladándolos a *F*. En la lista siguiente, sus razonamientos están en cursiva.

- **Dimensionalidad:** *F* es tridimensional. *Si V es tridimensional entonces cada tren de ondas viene descrito por su orientación (dada por dos ángulos, azimut y elevación) y por su frecuencia. La frecuencia en Y pasa a ser el radio de F, mientras que los dos ángulos de V siguen siendo ángulos en F. Si V fuera una esfera n-dimensional, F tendría también n dimensiones.*
- **Localización:** Los puntos que están próximos en *F* pueden intercambiar energía de formas predecibles y coherentes, mientras que los puntos que están alejados no pueden hacerlo. *Dos trenes de ondas de V que sean semejantes, por lo que a dirección y frecuencia se refiere, están en fase gran parte de su longitud y los efectos no lineales de agrupamiento serán más o menos los mismos un ciclo tras otro a lo largo de toda la longitud. Por otro lado, las configuraciones de onda distantes cuyas crestas y valles no sean correlativas perderán en unas y ganarán en otros y, en general, serán para las otras como simples golpes al azar.*
- **Velocidad de interacción:** Hay una velocidad característica para cada punto de *F*. Los puntos que están más alejados del centro de *F* interaccionan con mayor rapidez que los que están más cerca. *Una interacción es una transferencia de energía no aleatoria de una configuración de onda a otra. Un ciclo es la porción menor de un tren de onda que se repite. Un efecto que se repita de manera semejante en*

todos los ciclos puede tener un efecto constante sobre todo el tren de ondas. Los efectos que se producen en V se propagan a la velocidad del sonido, de forma que todo un ciclo puede quedar afectado en el tiempo que necesita el sonido para atravesarlo (este tiempo es el período de la onda). Las partes más externas de F corresponden a las configuraciones de onda con frecuencias más elevadas y, por lo tanto, a las mayores velocidades de interacción.

- **Principio de incertidumbre:** La energía de un punto de *F* no se puede determinar con precisión en un espacio de tiempo corto. La exactitud que se puede conseguir crece linealmente con la duración de la medida. *La energía de un punto de F es la energía total de un tren de ondas específico que ocupa todo el volumen V. Como ninguna señal de V puede desplazarse a mayor velocidad que la del sonido, hallar la energía total de un tren de ondas supondría esperar a que llegaran todas las señales desde todos los lugares de V, un tiempo mucho mayor que el tiempo básico de interacción. En un tiempo reducido, el total es necesariamente superior a un volumen proporcionalmente menor. Como el observador en F está distribuido por todo V, lo que no queda definido es cuál es el volumen proporcionalmente menor y, en consecuencia, la medida es incierta. A medida que aumentan el tiempo y el volumen total, las sumas posibles convergen hacia la suma media y la incertidumbre disminuye.*

- **Superposición de estados:** Parece como si la mayor parte de las interacciones que tienen lugar en *F* fueran la suma de todas las formas posibles en que se puede producir la interacción. *Cuando se produce una interacción entre dos trenes de ondas que están próximos, en un principio es ciclo por ciclo, ya que la información de partes lejanas de los trenes de ondas van llegando solamente a la velocidad del sonido. Cada uno de los ciclos contiene una pequeña parte de energía del tren de ondas en cuestión y una gran parte de energía de muchas otras ondas de distinta frecuencia y*

orientación que atraviesan el mismo volumen. A lo largo de todo el tren de ondas, el «ruido de fondo» será distinto de un ciclo al siguiente, de forma que la interacción será ligeramente distinta en cada uno de los ciclos. Cuando se ha reunido la información de todo el tren de ondas, la interacción total se puede interpretar como la suma de las interacciones de todos los ciclos. En ocasiones, la energía se transfiere a un ciclo de una forma y de la opuesta a otro ciclo más alejado, así que las energías se pueden sumar o cancelar.

Estas y otras propiedades del mundo *F* reflejan las características fundamentales de la mecánica cuántica, pero son solamente consecuencia de examinar una situación prosaica de una forma que se sale de lo normal. Existen algunas diferencias. La superposición de estados es estadística, no una suma perfecta de todas las posibilidades, como sucede en la mecánica cuántica tradicional. Esto supone sólo una ligera diferencia si *V* es muy grande, pero la cantidad de «ruido» es mínima en medidas que podrían ser útiles para distinguir el mecanismo de *F* de otras explicaciones de la mecánica cuántica.

El modelo, tal y como se ha presentado, no exhibe los efectos de la relatividad especial de ninguna forma evidente, lo que supone un serio defecto si esperamos enfrentarlo con una descripción de nuestro mundo. Hay algo equivocado en la forma de tratar el tiempo. Tiene una propiedad que imita los efectos temporales de un campo gravitatorio relativista general. El tiempo, cerca del centro de *F*, pasa con mayor lentitud que en los extremos, ya que las interacciones se basan en ondas de frecuencia más baja. En el mismo centro, el tiempo se detiene. El punto central de *F* tiene siempre la «densidad media de toda la esfera» y, por lo tanto, está realmente congelado en el tiempo. En la relatividad general, las regiones situadas alrededor de un cuerpo gravitatorio tienen una propiedad parecida: el tiempo pasa con mayor lentitud a medida que nos aproximamos. En las cercanías de masas muy densas (es decir, agujeros negros), el tiempo se detiene al llegar a una cierta distancia.

Es posible que algunas de las teorías más exóticas de la física moderna tengan alguna explicación plausible para este modelo. Aunque la energía se trasmite principalmente entre configuraciones de ondas de frecuencia y dirección semejantes (como los puntos adyacentes de F), la no linealidad del medio V permite que parte de la energía fluya sistemáticamente entre configuraciones de onda armónicamente relacionadas, por ejemplo, entre dos configuraciones que tengan la misma dirección pero la frecuencia de una sea el doble que la de la otra. Estas configuraciones del flujo de energía en F proporcionan otros «grados de libertad» además de los tres que proporcionan los puntos próximos. Cuando se observan a pequeña escala, se pueden interpretar como dimensiones adicionales (la energía se puede mover por aquí, por allí y por allí, y *también* por allí y por allí...). Como una vuelta completa, de un armónico a otro, cubre todo el espacio disponible en menos pasos que moviéndose a lo largo de configuraciones de onda adyacentes, parecerá que estas tres dimensiones adicionales tienen un alcance mucho mayor que las tres básicas. Cuanta más energía participa, se activan más armónicos y crece la dimensionalidad. En la mayor parte de las teorías físicas de la actualidad se «ata corto» a las dimensiones adicionales para dar una explicación geométrica a las fuerzas básicas. Es corriente hablar de diez y de once dimensiones, y las nuevas fuerzas que sugieren algunas teorías pueden dar lugar a más. Si resulta que es correcto algo semejante a la explicación F de una dimensionalidad aparentemente mayor, sería algo muy ventajoso. A gran escala, las «dimensiones» armónicas son vínculos reales entre regiones del espacio que se encuentran alejadas y, si se utilizaran adecuadamente, permitirían la comunicación instantánea y el desplazamiento a través de grandes distancias.

ONDAS GRANDES

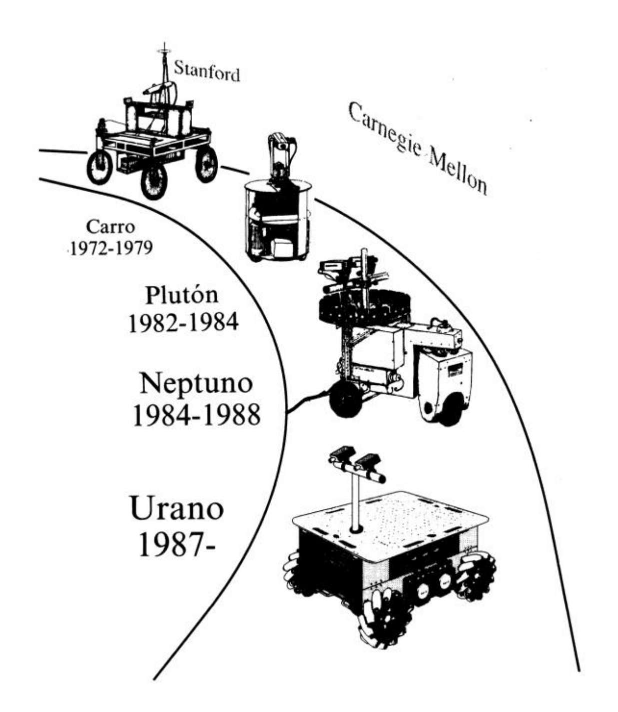
Olvidemos ahora todas las posibles implicaciones que supondría que F fuera un modelo adecuado para la mecánica

cuántica y consideremos nuestro universo a gran escala. Nuestro universo está impregnado por un fondo de radiación de microondas de una longitud de onda de un milímetro aproximadamente, longitud que aumenta lentamente a medida que el universo se expande y enfría. Influye sobre las nubes de materia (y, a su vez, ellas también lo afectan) del espacio interestelar, y, en consecuencia, interacciona consigo mismo de forma no lineal. Si hacemos una transformada de Fourier espacial, aplicada a todo el universo, de esta radiación (es decir, si consideramos que *nuestro* mundo es *V*), terminaremos con un espacio *F* que tendrá propiedades muy semejantes a las que aparecían en la sección anterior. Pero la expansión del universo añade un rasgo peculiar nuevo. A medida que nuestro universo se va expandiendo, las longitudes de onda de la radiación de fondo aumentan. A medida que estas longitudes de onda van creciendo, desciende la velocidad relativa del paso del tiempo en el mundo *F*. Si hubiera habitantes en F, estarían en el lugar idóneo para practicar la estrategia que propuso Freeman Dyson: «Vive eternamente moviéndote cada vez más despacio a medida que el mundo se enfría.» Ahora se tendrían que mover muy lentamente, ya que las interacciones de sus partículas más rápidas duran varias billonésimas de segundo. En el pasado, poco después del *big bang*, cuando el universo era denso y estaba caliente, el mundo *F* habría sido un lugar animado y bullicioso, que se movería a una velocidad millones o miles de millones de veces mayor. En los primeros momentos del universo, la velocidad sería astronómica.

El primer microsegundo después del *big bang* pudo representar siglos y siglos de tiempo subjetivo en *F*. Acaso el tiempo suficiente para que evolucionara la inteligencia, se diera cuenta de la situación y sembrara la vida, menor pero al final más rápida, en el espacio *V*. Aunque a gran escala *F y V son* lo mismo, sería extremadamente difícil que uno de ellos manipulara al otro, incluso que se produjera la comunicación entre los dos. Cualquier suceso local que tuviera lugar en uno de los dos espacios sería indetectable en el otro. Solamente funcionarían los

proyectos masivos, que abarcaran todo el universo, y precisarían enormes cantidades de tiempo debido a que en los dos universos existen límites para la velocidad. Queda descartada la interacción del tiempo real entre *V y F*. Sin embargo, esos proyectos podrían afectar a muchos puntos del otro espacio con la misma facilidad (a veces con mayor facilidad) que a uno solo, y podrían aparecer en ellos como «milagros» que violaban la entropía. Si yo viviera en *F y* deseara visitar *V* diseñaría uno de estos milagros y condensaría en *V* un robot que me sustituyera. Luego, por medio de otro milagro, podría leer los recuerdos del robot en una forma accesible de *F*.

La transformada de Fourier que convierte *V* en *F* es idéntica, excepto por un signo menos, a la transformada inversa. Si se proporcionan solamente las dos descripciones, no quedaría nada claro cuál era el mundo «original». En realidad, la transformada de Fourier no es más que otra de las infinitas «transformadas ortogonales» que tienen las mismas propiedades básicas. Todas ellas pueden aplicarse a una descripción de un volumen y transformarla en una descripción diferente que contenga la misma información, pero todos los puntos originales se habrían diseminado por el resultado. Esto nos sugiere la posibilidad de que exista una infinidad de universos cada uno de los cuales sería una combinación diferente del material inicial. El comportamiento de todos ellos se ajustaría a la mecánica cuántica, pero cada uno tendría sus propias leyes físicas, independientes de las de los otros universos con los que compartir el espacio. No sé qué hacer con esta idea.



Compañeros robots
Algunos de los colaboradores mecánicos del autor y los años que duró la asociación.

AGRADECIMIENTOS Y CRÉDITOS DE LAS ILUSTRACIONES

Este libro tiene sus raíces en mi infancia. No tengo tan buena memoria como para poder identificarlos individualmente, pero quiero expresar mi agradecimiento a los autores de ciencia y de ciencia-ficción, profesores, bibliotecarios, organizadores de reuniones científicas y amigos que, a lo largo de cuatro décadas, me han ayudado a conformar mi mundo mental. Mi memoria, sin embargo, *sí* es lo suficientemente buena como para permitirme recordar que mis hermanas pequeñas, Elizabeth y Alice, han sido durante muchos años las pacientes cajas de resonancia de mis largas y emocionadas especulaciones.

A finales de 1971, cuando llegué al Laboratorio de Inteligencia Artificial de Stanford (SAIL) como estudiante graduado, se estaba concluyendo un interesante debate que había puesto sobre el tapete un proyecto elaborado por Dick Fredericksen y aparecido en unos boletines informativos que publicaba él mismo, *A Word in Edgewise*. A lo largo de varios artículos, había desarrollado la idea de conseguir la inmortalidad, y mucho más, sustituyendo el sistema nervioso, trozo a trozo, por un equivalente artificial mucho más duradero. Me pareció interesantísimo, porque la idea se me había ocurrido a mí años antes, cuando todavía estaba en el colegio, pero nadie me había hecho mucho caso. En el SAIL, los que se habían tomado en serio el proyecto de Fredericksen, habían adoptado posturas extremas y enfrentadas. Bruce Baumgart capitaneaba a los que estaban a favor, mientras que Larry Tessler lo encontraba deshumanizador. Mis pensamientos sobre el futuro de las máquinas inteligentes cristalizaron en las discusiones que mantuve con Rod Brooks, Bruce Bullock, Mike Farmwald, Bob For-

ward, Don Gennery, Erik Gilbert, Bill Gosper, David Grossman, Brian Karvey, Marc Le Brun, Robert Maas, John McCarthy, Ed Mcguire, Dave Poole, Jeff Rubin, Clem Smith, Russ Taylor, Lowell Wood y algunos otros. En 1975 escribí un ensayo sobre el tema que, con el paso de los años, se convirtió en varios artículos y, finalmente, en este libro.

Cuando, en 1980, me trasladé a la Carnegie Mellon University, las discusiones continuaban. Tengo el placer de agradecer los enriquecedores intercambios de ideas que mantuve con Mike Blackwell, Kevin Dowling, Alberto Elfes, Larry Matthies, Pat Muir, Gregg Podnar, Olin Shivers y Richard Wallace. También me gustaría dar las gracias a la sección administrativa del Instituto de Robótica, en especial a Raj Reddy y a Takeo Kanade, por mantener un entorno que me permitió dedicarme a los objetivos a largo plazo. Quiero, asimismo, dar las gracias al Departamento de Investigación Naval y al director de mi programa, Alan Meyrowitz, por proporcionarme los fondos con los que he desarrollado mis investigaciones básicas desde 1981.

Aunque ya en 1975 tenía la vaga intención de desarrollar mis ideas y convertirlas en un libro, no empecé a trabajar en el manuscrito hasta 1985. Debido a una asombrosa coincidencia, dos semanas después de comenzar el proyecto recibí una carta de Howard Boyer, que acababa de tomar posesión de su puesto de Editor de Ciencia y Medicina de Harvard University Press, en la que me invitaba a escribir un libro sobre el mismo tema. En los tres años siguientes, Howard hizo posible que el libro siguiera adelante venciendo una serie de penosos obstáculos. Le estoy profundamente agradecido por su interés y su intuición.

El primer obstáculo fue elaborar un esquema detallado que mereciera la aprobación de Harvard Press, y se superó con la ayuda de los exhaustivos análisis que escribió Vernor Vinge. Los numerosos borradores que siguieron fueron muy mejorados por los asesores de Press, con sus comentarios alentadores, duros e informativos. Quiero expresar mi agradecimiento a Rod Brooks, Richard Dawkins, Kee Dewdney, Bruce Donald,

John Dowling, Bob Forward, John McCarthy, Pamela McCorduck y otros cuya identidad desconozco.

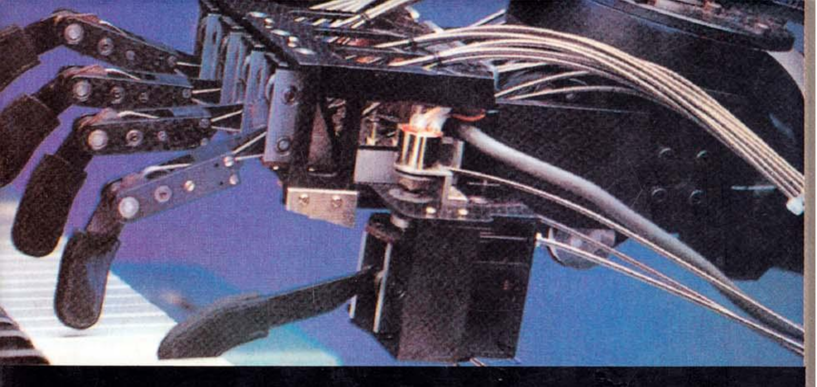
Las mayores reformas del libro tuvieron lugar cuando se puso en manos de la editora el manuscrito, Susan Wallace. Susan reorganizó el texto y convirtió en un conjunto coherente una confusa colección de ideas. Con ello puso las bases para hacer una nueva versión de la obra que, en cuanto a calidad, era completamente distinta a la anterior.

Con las excepciones que mencionaré después, el libro lo he elaborado yo mismo con un Macintosh II de Apple Computer. Los programas que he usado han sido *Cricket Draw*, de Cricket Software, y *SuperPaint,* de Silicon Beach Software, con algunos préstamos ocasionales de transcripciones digitales de colecciones de *clips:* los paquetes *McPic!* de Magnum Software y los *Click Art* de T/Maker Graphics. Con el programa *Thun derscan* y *hardware* de Thunderware se examinaron algunos dibujos hechos a mano y fotografías. Las primeras versiones de muchos de los dibujos se hicieron con pequeños ordenadores Macintosh y con los programas *MacPaint, Mac Draw, Full Paint* y *Mac Draft*

Mike Blackwell volvió a dibujar «Inteligencia sobre la Tierra» con *Cricket Draw* a partir de mi versión original con *Mac Draw.* «La retina» la dibujó Mary Jo Dowling utilizando el *Adobe Illustrator* de Adobe Systems. Lo hizo a partir de una ilustración que apareció en *The Retina,* de John Dowling (sin ningún parentesco con ella). Los iconos de los mecanismos y circuitos integrados de «Un siglo de cálculos» los dibujó Gregg Podnar con un *MacPaint.* «Un robot arbusto» se elaboró con un programa que yo mismo escribí, ayudado por Mike Blackwell, con lenguaje Apple Computer's *MPW C* y un ordenador Macintosh II. Tiene un cuarto de millón de segmentos de línea y tardé diez horas en hacerlo. El dibujo de los «Marcianos egoístas» lo hizo Kimberlee Faught con un *Cricket Draw.* Los dibujos del Carro, Plutón y Neptuno que aparecen en «Compañeros robots» son transcripciones digitales y reproducciones retocadas de dibujos hechos a lápiz por Bill Nee. El dibujo de

Urano es una transcripción digital de una foto que retocamos Gregg Podnar y yo.

Las siguientes organizaciones me proporcionaron fotografías y me concedieron permiso para reproducirlas: «La máquina que anda», con permiso de Odetics, Inc.; «Tres dedos», con permiso de David Lampe, MIT; «Navegación autónoma», con permiso de Denning Mobile Robotics, Inc.; «Cómo se encuentran objetos», con permiso de SRI International; «La retina», con permiso de John Dowling; «ENIAC», © Smithsonian Institution; «Gafas mágicas», con permiso de United Technologies/Hamilton Standard; «Un agente robot», con permiso de Naval Ocean Systems Center; «Propiedades inmobiliarias irreales. La carretera a Punto Reyes», © 1986 Pixar.



Mientras sigue la discusión filosófica sobre si las máquinas pueden o no pensar, los expertos en inteligencia artificial y en robótica ya construyen máquinas inteligentes que están transformando nuestro entorno social, y esto es sólo el principio. En este libro, Hans Moravec explica la historia de la robótica y de la inteligencia artificial, describe el estado actual de estas ciencias, y predice un verosímil futuro en el que los robots inteligentes superarán los mejores logros de la mente humana y nos suplantarán en la implacable carrerra de la evolución.

Hans Moravec es uno de los investigadores de más renombre mundial en el campo de la robótica, y en la actualidad dirige el laboratorio de robots móviles de la Universidad Carnegie Mellon.